UNIVERSITY OF CALIFORNIA SAN DIEGO

Towards Improved Force-Field Accuracy for Calculation of Binding Thermodynamics

A dissertation submitted in partial satisfaction of the requirements for the degree Doctor of Philosophy

in

Chemistry

by

Sophie Michelle Kantonen

Committee in charge:

Professor Michael K. Gilson, Chair
Professor J. Andrew McCammon, Co-Chair
Professor Ruben Abagyan
Professor Michael Galperin
Professor Partho Ghosh

2020

This Dissertation of Sophie Michelle Kantonen is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

Co-Chair

Chair

University of California San Diego

2020

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES

ACKNOWLEDGMENTS

First and foremost I must acknowledge Dr. Michael Gilson for his support and guidance while I have been at UCSD. He took a chance on letting an incoming graduate student initiate experimental work in his laboratory, and I have been grateful for this opportunity ever since. I have probably learned more in the Gilson laboratory than any other scholarly period in my life, and Dr. Gilson always encouraged and facilitated curiosity and exploration (the best parts of science!). He truly is one of the best, not just in terms of world-class academic research, but also in the fostering of students who have the privilege to work in his lab, and for his sheer enjoyment of ethical and rigorous science that is always contagious. He is always extremely understanding and personable, and is first and foremost, a very good person. Thank you for the hours and hours of discussion; it truly has made my life better.

I'd also like to acknowledge my family; I'd like to particularly thank my mother and father, Lori and Andrew Kantonen, for a lifetime of unwavering support and encouragement. I can say with indefatigable certainty that I would be in a much worse place if not for their love and support. I'd also like to thank my two younger sisters, Elli and Lilja, who have also been a constant source of love and support in my life. Thanks for getting me through all of it, guys.

Lastly, I'd like to acknowledge all of the friends and co-workers I have met during my time at UCSD. Niel Henriksen and Dave Slochower took a lot of time to help me learn a lot not just about the science, but about the programming and key ideas behind the science. They also provided engaging scientific discussions in the lab, and I learned a lot just by talking with them. I'd also like to thank Hari Muddana, who took time to work with

me on the force-field development project. I'll miss the evenings walking to In n' Out and talking quantum chemistry.

Chapter 2, in full, is a reprint of the material as it appears in "Data-Driven Mapping of Gas-Phase Quantum Calculations to General Force Field Lennard-Jones Parameters", SM Kantonen**,** H Muddana, M Schauperl, NM Henriksen, LP Wang, MK Gilson. Journal of Chemical Theory and Computation (2019). The dissertation author was primary investigator and author of this publication.

Chapter 4, in full, is a reprint of the material as it appears in "Evaluation and minimization of uncertainty in ITC binding measurements: heat error, concentration error, saturation, and stoichiometry", in SM Kantonen, NM Henriksen, MK Gilson. Biochimica et Biophysica Acta (BBA)-General Subjects (2017). The dissertation author was primary investigator and author of this publication.

Chapter 5, in full, is a reprint of the material as it appears in "Accounting for apparent deviations between calorimetric and van't Hoff enthalpies", SM Kantonen, NM Henriksen, MK Gilson. Biochimica et Biophysica Acta (BBA)-General Subjects (2018). The dissertation author was primary investigator and author of this publication.

VITA

2013	Bachelor of Science, Wright State University
2014	Master of Science, University of California San Diego
2014-2020	Research assistant, University of California San Diego
2020	Doctor of Philosophy, University of California San Diego

ABSTRACT OF THE DISSERTATION

Towards Improved Force-Field Accuracy for Calculation of Binding Thermodynamics

by

Sophie Michelle Kantonen
Doctor of Philosophy in Chemistry
University of California San Diego
Professor Michael K. Gilson, Chair
Professor J. Andrew McCammon, Co-Chair

Molecular dynamics simulations have revolutionized chemistry by allowing cheap and fast *in silico* analyses of numerous systems of interest. Despite various advancements, many applications still require costly experimental input in order to guide direction, as the results from simulations are not yet accurate enough to rely on alone. This is particularly problematic in drug design, where accurate binding affinity measurements could greatly improve the ability to discover drugs. In this thesis, I discuss a philosophy behind improving force-fields, the functions which provide the configurational energies in simulations, for binding calculations. I show a new approach to non-bonded force-field parameterization, which reduces the number of parameters used and also preserves the chemical uniqueness of each atom in a molecule.

Secondly, I discuss a synthesis pathway towards generating novel host molecules for parameterization of force-fields built for binding calculations. Lastly, I present a novel and systematic analysis of experimental uncertainties in isothermal titration calorimetry data, to establish a clearer foundation for their use in force field parameterization. Taken together, these efforts contribute to an overall goal of developing force-fields that yield more accurate binding calculations.

1 INTRODUCTION

1.1 SYNOPSIS OF INTRODUCTION

In this thesis, a general philosophy for improving force-fields for binding affinity calculations is explored. This Introduction, Section 1, presents concepts relevant to experimental binding measurements and classical molecular dynamics simulations. Thus, Sections 1.2 and 1.3 give a general background on computational chemistry and classical simulations. These sections are meant to provide condensed information relevant to the overall topic. Section 1.4 gives an overview of commonly used methods binding thermodynamics are calculated and provides accompanying generalized equations. Section 1.5 gives a relevant discussion of force-field development, which is expanded upon in Chapter 2. This section also provides some background on the algorithm ForceBalance utilizes. Section 1.6 gives some background on host-guest chemistry, including why they are relevant to force-field development and fit in with the overall philosophy of the thesis. Details on this subject are expanded in Chapter 3. Section 1.7 details information on isothermal titration calorimetry, including a brief discussion into how it works and potential pitfalls, which are the subjects of Chapters 4 and 5. Finally, Section 1.8 provides a structure for the rest of the thesis and provides a statement of the overall individual goals each chapter was meant to play.

1.2 COMPUTATIONAL CHEMISTRY

Computers have revolutionized nearly every field they have been applied in, and this is particularly true in the field of chemistry. By opening up the possibility of doing many calculations in a short period of time, computers added both classical[1,2] and quantum mechanical simulations[3,4] to the toolbox of chemistry, allowing in depth

exploration of microscopic events at the molecular level. Molecular dynamics treats particles classically[1] and can provide information about the thermodynamics of a system of particles[2], while quantum mechanical simulations provide approximations of the electronic structure of atoms of molecules[4,5]. Here, we will focus on molecular dynamics, with a specific application towards measuring noncovalent binding affinities, although we will also touch on ways in which quantum chemistry is applicable to this endeavor as well.

Over the last five decades, during which computational chemistry has risen to a prominent field, molecular dynamics (MD) simulations[6–8] have been applied to a diverse range of research problems, including liquid structure[2,9–11], drug design[12,13], protein structure prediction[14,15], and materials research[16,17]. By utilizing an applied knowledge of statistical mechanics[18], molecular dynamics simulations can extract useful thermodynamic information regarding a given system. Remarkably, by setting a few key conditions, such as how the pressure and temperature are maintained during the simulation, along with parameters which define the interatomic forces, a molecular dynamics simulation can reproduce Boltzmann statistics for a given system, giving results which largely resemble what would be observed in the real world. This ability allows computational chemists to make predictions on how a given system will behave before any wet lab work needs done. This is particularly useful in computer aided drug design (CADD), where a large number of compounds can be screened *in silico* before any synthesis or assays are done[13]. However, in order for any useful information to be extracted from any molecular dynamics simulation, a potential function-- the function that determines the energy of each configuration of the system—is required. Much effort has been exerted towards creating accurate, computationally efficient potential functions, or

force fields[19–26], yet current results still seem to indicate[27–31] that these potential functions are not yet finely tuned enough to avoid requiring extensive experimental trial and error in the design of promising candidate drug molecules.

1.3 TECHNOLOGY OF MOLECULAR DYNAMICS SIMULATIONS

As discussed above, one of a computational chemist's most useful tools is the molecular dynamics simulation. In these simulations, particles are treated classically, and the energy and atomic forces at each configuration is given by a potential function. MD simulations are moved in incremental time steps by integrating Newton's Laws of Motion, given the interatomic forces and atomic masses. The potential function in turn contains a set of numerical parameters, which have been tuned to replicate reference data drawn from experiment or quantum mechanical calculations[19,21,24].

Given the functional form of a force-field and a suitable set of force-field parameters, a simulation can be carried out by selecting the time step, or how often the system configuration is updated, and selecting the ensemble the simulation should be run in; e.g., constant pressure (NPT) or constant volume (NVT). Simulations will be run at a target temperature, for example 310 Kelvin if one wishes to run a system at body temperature to study protein behavior, which is enforced by a selected thermostat[32,33]. Similarly, a barostat can be employed to maintain a constant pressure. These constraints give control over thermodynamic variables during the simulation and thus make it possible to extract useful information about the macroscopic thermodynamic behavior of a given system. This is essential to understand to accurately obtain relevant information about a system with respect to experiment. For example, a typical binding measurement in our laboratory is run at constant pressure; running a NPT simulation to replicate that binding event would be more suitable, as a NVT simulation might

experience wild pressure changes in an improperly equilibrated simulation. These simulation controls thus allow the user to replicate key experimental conditions.

There are additional key assumptions and constraints that are made on the system to improve accuracy and speed. One is the construction of periodic boundary conditions; which mimic an infinite system by treating the simulation box as periodic in space[34]. For example, if a molecule exits the left side of the simulation box, it will reenter from the right. Additionally, atoms interact with each other across the cell boundaries. This is a relatively cheap and effective method of avoiding artifacts at the simulation boundary.

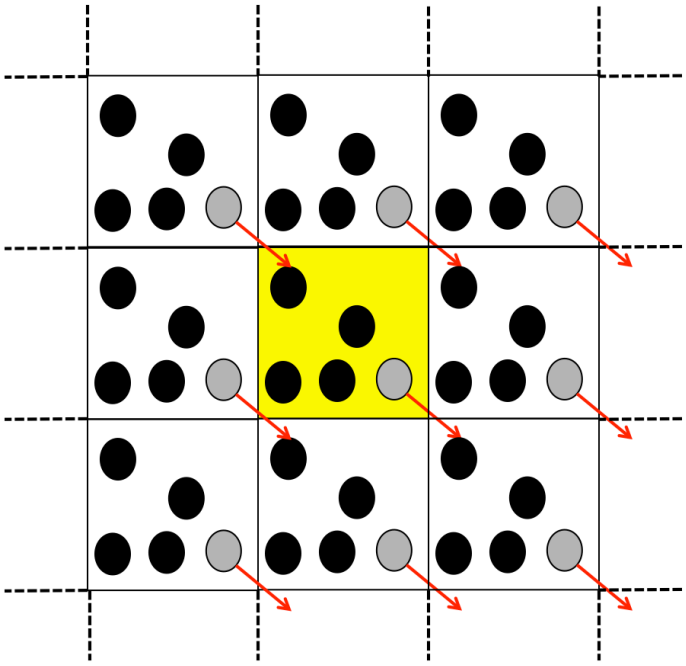


Figure 1.1: A two-dimensional diagram schematic of periodic boundary conditions. Particles in the highlighted central unit cell are actually simulated, while copies are built around it. Velocities are conserved as particles pass between unit cells, as illustrated by the grey particle.

A second approximation is the treatment of cutoffs in the non-bonded terms. For electrostatics, cutoffs can be avoided by utilizing Particle Mesh Ewald interactions (PME)[35]. In PME, electrostatic interactions are classified into short and long range interactions. The short range calculations are done normally, taking the electrostatics as the term seen in the potential function. After a pre-set distance, the long-range electrostatics are calculated as Ewald lattice sums, efficiently calculated making use of Fast Fourier Transform. This allows the simulation to be run much faster while not sacrificing much in terms of accuracy. Similarly, non-bonded terms are typically cutoff past a certain distance to hasten simulation time.

1.4 POTENTIAL FUNCTIONS

Potential functions are critical components of molecular dynamics simulations, and small inaccuracies can have large consequences on the behavior of the simulated system. Perhaps the most commonly used form for the potential function can be written as:

$$
\begin{aligned}
E(\boldsymbol{r}) = & \sum_{\text{bonds}} K_{\text{r}}\left(r - r_{\text{eq}}\right)^2 + \sum_{\text{angles}} K_{\theta}\left(\theta - \theta_{\text{eq}}\right)^2 \\
& + \sum_{\text{torsions}} \sum_{n=0}^{5} C_n(\cos(\psi))^n + \sum_{\text{impropers}} k_{\text{d}}\left(1 + \cos\left(n_{\text{d}}\omega - \omega_{\text{d}}\right)\right) \\
& + \sum_{i>j}^{N}\left[\left(4\varepsilon_{ij}\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right) + \frac{1}{4\pi\varepsilon_0}\frac{q_i q_j}{r_{ij}}\right]
\end{aligned}
\tag{1}
$$

Here, E(***r***) is the total potential energy as a function of the configuration of the system; i.e., the atomic coordinates ***r***. The first three sums represent the potential energy from all of the chemical bonds, angles, and dihedrals, with K_r and K_q being spring constants with respective equilibrium distances of bonds and angles r_{eq} and q_{eq}, and C_n being the amplitude of the potential energy function of angle ψ, with n nodes. An additional improper term, controlled by a stiffness factor, k_d, governs the angle ω's movement away from equilibrium angle ω_d, giving some control to the structure of the molecule. The last two sums represent non-bonded potential energies: the Lennard-Jones model of van der Waals interactions and Coulombic interactions. Parameters A_{ij}, B_{ij} correspond to the repuslive and attractive parts of the Lennard-Jones model, respectively and q_i and q_j being partial charges for atoms i and j, with ε_0 being vacuum permittivity. While this form of the potential function is very common, other forms account explicitly for additional physical effects such as electronic polarizability and/or modify the form of the non-bonded potentials. The common thread all of these functions share is that they require some sort of input parameters.

It is worth considering the origin of the Lennard-Jones terms which appear in Eq. 1, which are of particular interest in relation to the noncovalent interactions that drive the formation of supramolecular complexes. As mentioned, there is an attractive and repulsive component that make up the total Lennard-Jones interaction, with potentials that go as $1/r^6$ and $1/r^{12}$, respectively.

The attractive portion of the Lennard-Jones is due to dispersion forces[36]; the instantaneous attraction between atoms due to the wave-like nature of the electron density. The attractive part of the Lennard-Jones model itself is derived from another

model of the dispersion forces called the London dispersion model[37]. In London's model, the atom is modeled as a fluctuating dipole, with a characteristic frequency and polarizability. London showed that the dispersion interaction could be written as:

$$E_{AB}^{disp} \approx -\frac{3}{2}\frac{I_A I_B}{I_A + I_B}\frac{\alpha_A \alpha_B}{r^6} \qquad (2)$$

Here I_A and I_B are the first ionization potentials of atoms A and B, respectively, α_a and α_b are polarizabilities of atoms A and B, respectively, r is the interatomic distance, and E_{AB} is the potential due to dispersion between atoms A and B. Higher moments can be considered, expanding out to fluctating quadrapole or octapoles, but in general the term is simply taken as Eq. 2.

The repulsive part of the Lennard-Jones model comes from exchange repulsion[36], a purely quantum mechanical effect that arises due to the fermionic nature of electrons. The exchange repulsion has been shown to be well-approximated by the overlap of electron densities, and such information is readily obtainable from relatively straightforward and simple quantum mechanical calculations. Similarly, the ionization potentials and polarizabilities can be estimated from the electron density. The use of such data to drive parameterization is discussed in detail in Chapter 2.

1.5 Using Molecular Dynamics to Compute Noncovalent Binding Affinities

On their own, molecular dynamics simulations provide rather limited information about the macroscopic system. There are multiple ways to extract a great deal of useful information from simulations, such as structural information[10,14,15], kinetics[38–40], and thermodynamics[18,41,42]. Here, I will focus on using MD to extract information about

binding affinities. In order to accomplish this, a number of methods exist which can provide information about the relative free energy of the system. By simulating a given protein-ligand system, information about the change in free energy upon complexation can be given[42,43], which can be useful information in determining whether a compound will be a useful drug. The relationship between the binding constant, K_B, and the free energy change upon binding can be written as:

$$K_B = e^{\frac{-\Delta G^o}{RT}} \qquad (3)$$

Here, K_B is the forward equilibrium constant of complexation, ΔG^o is the free energy of binding, R is the gas constant, and T is the temperature at which binding occurs. Of course, having an accurate estimate of the affinity is of great use for screening drugs *in silico,* so efforts in obtaining ΔG^o have been of great interest to those in drug development.

One widely use method of free energy calculation is thermodynamic integration (TI)[44]. Here, the free energy is governed by a variable lambda, which represents a scaling between an initial and final state of interest. Integrating the partial derivative of the free energy with respect to the lambda variable over a range of lambdas utilized during the series of simulations provides the difference in free energy between the two states, which can provide the binding free energy when performed between the bound and unbound states of a protein-ligand system:

$$\Delta G = \int_0^1 \langle \frac{dU(\lambda, \boldsymbol{r})}{d\lambda} \rangle_\lambda \, d\lambda \qquad (4)$$

Here ΔG is the Gibbs free energy difference between λ=0 and λ=1, and U(λ,**r**) is the potential energy of the system at a given lambda and set of coordinates **r**. Angle brackets indicate Boltzmann averages, in this case the average value over number of frames simulated. While TI is simple and straightforward, it has several potential drawbacks. First, if the curvature of dU/dλ is too great, it can lead to large errors in the final calculated free energy change, due to bias introduced from the discrete values of lambda. This can arise particularly when dealing with initial and final states very far apart from one another in terms of energy landscapes. Changing too much too fast when scaling between the two states in lambda can lead to failures in the method. Second, is the need to calculate derivatives at each simulation, which can be not as straightforward as merely taking an energy or other simple output from simulation.

Another method of free energy calculation is free energy perturbation (FEP). This method involves what is termed the Zwanzig equation[45]:

$$\Delta G(A \rightarrow B) = G_B - G_A = -k_B T ln \langle exp(\frac{U_B(\boldsymbol{r}) - U_A(\boldsymbol{r})}{k_B T}) \rangle_A \qquad (5)$$

Here, ΔG is the free energy difference between states A and B, k_B is the Boltzmann constant, U_A and U_B are energies of states A and B, respectively. As opposed to TI, which computes derivatives of the energy with respect to some parameter, which scales between the two states, FEP takes the energies between two different states to estimate the difference in free energy. The key requirement of this method is that the states overlap; that is to say, the two systems must at least partially occupy the same phase space at the temperature the simulation is being run at. This can be difficult in terms of going straight from bound to unbound in a protein-ligand complex, for example. This can be fixed by appropriately using a number of windows between the bound and unbound

state. By measuring the free energy between small changes, overlap can be better achieved. If the differences are too large, the calculation will never converge. Another related method is the Bennett Acceptance Ratio (BAR)[46], which can generally be written as:

$$e^{-\beta(\Delta G-C)} = \frac{\langle f(\beta(U_B(\boldsymbol{r}) - U_A(\boldsymbol{r}) - C))\rangle_A}{\langle f(\beta(U_A(\boldsymbol{r}) - U_B(\boldsymbol{r}) + C))\rangle_B} \quad (6)$$

Here, f(x) is the Fermi-Dirac distribution, U_A and U_B represent potential energies calculated using partition functions from system A or B, respectively, while angular brackets again indicate averaging over configurations at state A or state B; C represents an energy offset which is typically numerically solved in order to obtain ΔG. An extension of BAR by Shirts and Chodera showed that the method could be applied over multiple states, appropriately naming the method Mulistate Bennett Acceptance Ratio (MBAR)[47]. This is relevant to the problem of overlap, as BAR requires similar restraints as mentioned above in terms of overlap of energy densities of the states being examined. Thus, a reaction coordinate can be broken up over smaller increments and the overall free energy difference can be taken with MBAR.

While these methods provide a theoretical framework to calculate binding free energies, more specific algorithms are required to actually extract this information from simulations. One of the most straightforward methods, double decoupling[42], involves gradually turning off the interactions of the ligand from the protein (decoupling), then the solvent, in a free energy cycle that provides the free energy of binding when treated with one of the aforementioned free energy methods. Another method, Attach-Pull-Release (APR), this method computes the reversible work of linking the ligand to the protein by a

spring, then of extending the spring to pull the ligand out of the binding site, then of releasing the ligand to bulk solvent[28,43]. In order to pull the ligand out, the equilibrium length of the spring is increased in incremental windows, in a method referred to as umbrella sampling. Here, the position of the ligand is restrained of the spring, and the force (for TI) or energy (for MBAR or FEP) on the spring at each distance is calculated. The sum of the reversible work terms for the attach, pull, and release steps gives the binding free energy of the ligand. This method has an additional benefit, as the mean potential energies of the unbound and bound states can be saved to determine the binding enthalpy as well. This additional thermodynamic information provides insight into the mechanism of binding (i.e., enthalpic or entropic driven binding), and can give additional insight into the accuracy of the simulation in comparison with experimental binding enthalpies. This combined method is referred to in our laboratory as computational calorimetry[43], as we are provided the same information from the simulation as an experimentalist is provided from a binding calorimetry experiment.

1.6 FORCE-FIELD DEVELOPMENT

A number of methods have been developed to more adequately sample the configuration space of a given system, including umbrella sampling[44] (as previously mentioned), metadynamics[48–50], Accelerated MD[51,52], and others[53]. Additionally, hardware advances[54,55], such as application of GPU's[56,57], have further allowed MD simulations to adequately sample configuration space. While these advancements have led to overall improvement in the capabilities of MD, there are still often substantial inaccuracies in the calculations of binding thermodynamics. These inaccuracies are thought to exist in major part due to the force-field used[27,28,58,59], as there are a number of approximations made that limit the accuracy of what simulations can determine. Thus,

making improvements in force-fields with the goal of reproducing condensed phase properties has been a major research goal for the last two or three decades.

As mentioned above, the most straightforward and often used functional form (equation 1) makes key approximations about how the physics are represented which can lead to limitations in accuracy of simulations. One of the most prominent is the use of classical, atom-centered, point charges, where the electrostatic interactions are approximated as interactions among partial charges at the nuclear coordinates of the atoms. This neglects key, well-known phenomena such as charge penetration[60,61] and the presence of sigma holes[62,63] in distribution of electrons around a given atom. Additionally, the Lennard-Jones model represents a simplistic model of non-bonded interactions, simplified to improve speed of calculation, and gives potentials that may not accurately represent realistic interatomic potentials[64].

An additional difficult step of force-field development is the successful parameterization of a given force-field. There are many parameters necessary for the calculation of the potential energy, and these parameters possibly can be unique for every unique atom, bond, angle, and torsion in the molecule, leading to an explosion of different parameters to be correctly determined. Thus, when developing a force-field it is common to use atom-types, in which identical parameters are decided for a group of atoms, such as the oxygens found in carbonyl groups. This allows for more expedient and straightforward parameterization against reference experimental data; including too many atom-types requires a very large set of reference data. While atom-typing allows for more efficient optimization, it is often far from clear how best to classify atoms into types[65], and it is not clear that the types used in the most common force fields have been derived by any systematic method. Additionally, there is also a risk of being too

conservative with the number of atom types, as inaccuracies can result from grouping chemical diverse atoms into a single type. Taken together, atom-typing presents a serious problem in terms of developing force-fields, and attention has been paid to either systematically define atom types or get rid of them entirely.

Efforts to avoid atom-typing altogether have been bolstered by the Tkatchenko-Scheffler principle of generating non-bonded forces[66]. Jorgensen showed that electron densities of atoms-in-molecules could be used to inform parameterization of Lennard-Jones parameters[67]. Similarly, QUBEKit also generates Lennard-Jones parameters based on electron densities[23]. These methods circumvent atom-typing by utilizing on the fly quantum mechanical calculations of molecules prior to simulation, which are used to generate parameters, as opposed to reading in an array of pre-set parameters. Additionally, efforts have been made in the OpenFF collaboration to assign parameters in a more elegant way, based on direct chemical perception of the inputted molecular structure as opposed to a collection of bonded atom-types[65].

Once a set of force field parameters has been defined, they should, ideally, be systematically optimized in order to ensure maximum accuracy. While certain terms, such as bonded or angle terms, can be fairly easily determined from quantum mechanical calculations, non-bonded terms are harder to define with QM calculations, as values fitted to QM may not work well at yielding accurate condensed-phase properties[36,64,68]. As a consequence, it is often preferable to fit them to measured properties of condensed-phase systems[19,22,24,26,69]. A clear example of this is the Lennard-Jones term, where the repulsive LJ component of the parameters, A_{ij}, can greatly influence the density of a neat liquid, because A_{ij} essentially being the diameter , or size of the given atom it represents. Similarly, the attractive LJ parameters, B_{ij}, are

important determinants of the internal energy of liquids. Thus, one of the most successful force field parameterization efforts, the OPLS force-field, used simulations of neat liquids to systematically adjust parameters by comparing densities and heats of vaporization of neat liquids to experimental values[24]. This took a step to ensuring that the parameters were sound and, at least for liquid data, reproduced experimental properties. However, it does not appear that the the full combinatorial parameter space of current force fields was systematically explored in this pioneering work.

Indeed, although adjusting parameters against experimental reference data is a powerful approach, it can be difficult to adequately sample what ultimately becomes a high-dimensional parameter space; there may not be enough suitable reference data to cover the diversity of chemical space; and using simulations to compute a large number of experimental observables within an optimization loop is computationally costly.

A major improvement towards parameter optimization came with the use of sensitivity analysis[19,69], wherein the gradient of parameters with respect to a given thermodynamic variable is computed in order to best determine what steps to take in parameter tuning:

$$\frac{d}{d\gamma}\langle A\rangle_\gamma = \frac{1}{Q}\int \exp\big(-\beta(E+PV)\big)\left[\frac{\partial A}{\partial \gamma} + A\left(-\beta\frac{\partial E}{\partial \gamma}\right)\right] dr dV - \frac{1}{Q^2}\frac{dQ}{d\gamma}\int A\exp\big(-\beta(E+PV)\big)\, dr dV = \langle\frac{\partial A}{\partial \gamma}\rangle_\gamma - \beta\left(\langle A\frac{\partial E}{\partial \gamma}\rangle_\gamma - \langle A\rangle_\gamma\langle\frac{\partial E}{\partial \gamma}\rangle_\gamma\right) \quad (7)$$

Here, γ is a given force-field parameter, P is the pressure of the system, V is the volume of the system, and A is a generic thermodynamic property of the system. Ensemble averages of $dA/d\gamma$ can be calculated using the above relationship, which can be calculated by re-running the trajectory of a simulation. This principle can be extended to many different properties, such as the density and heat of vaporization of a liquid, or even binding free energies. This method has been successfully utilized to perform guided optimization on TIP3P water, using the ForceBalance suite of software. This software runs simulations, computes gradients, and updates parameters iteratively to search for the best optimized parameters. Our laboratory has also successfully utilized this method of sensitivity analysis with respect to binding free energies and enthalpies to guide optimization of a small set of force-field parameters.

Overall, a method to reduce the parameter space while maintaining the unique properties of each molecule would significantly improve efforts to develop force-fields. Based on this principle, and the pioneering work of others[23,66,67,70], I have developed a

method which both discards atom-types and maintains the uniqueness of each atom in a given molecule. Additionally, the parameters are generated using information from quantum mechanical calculations in tandem with experimental data. This method, called the Slater-Derived Lennard-Jones (SDLJ), is detailed in Chapter 2.

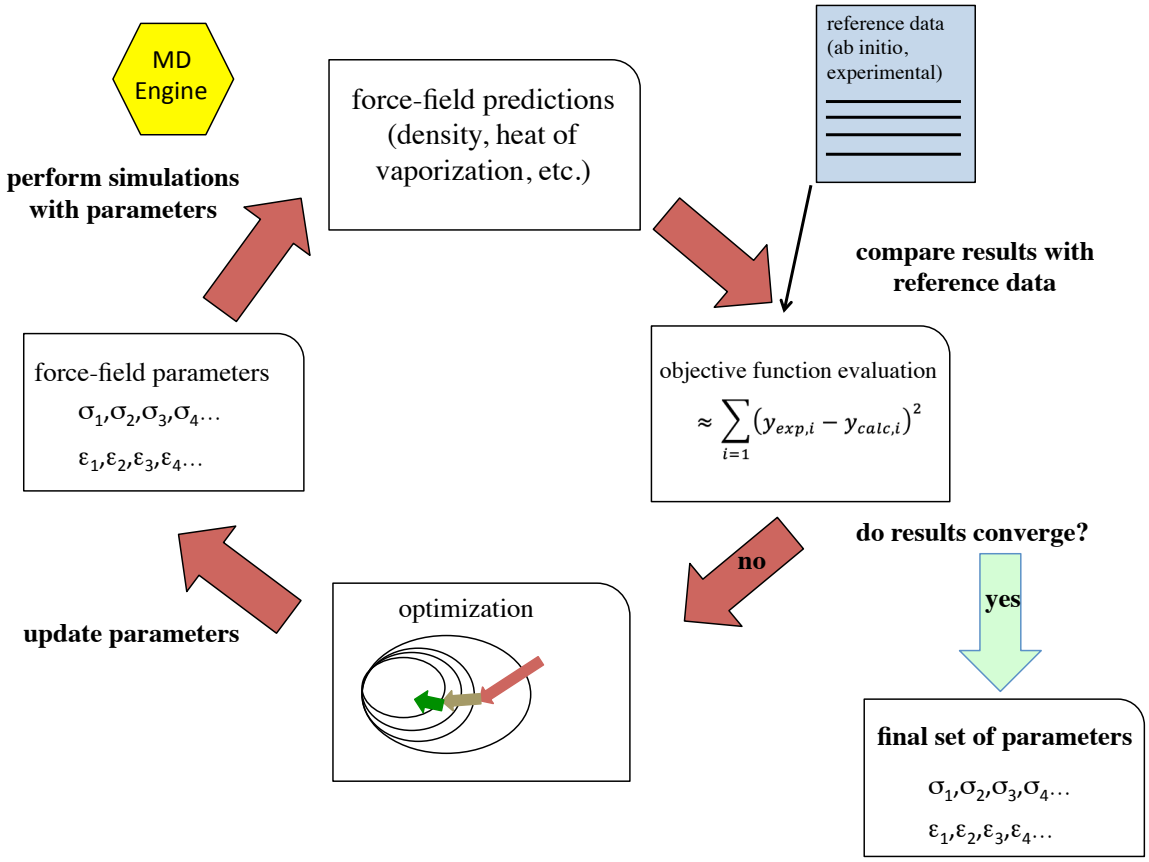


Figure 1.2: General workflow of force-field optimization utilized in particular by ForceBalance; starting at the yellow hexagon, a MD simulation makes a calculation, which is compared to experimental data. The gradient is computed via an objective function and the parameters are adjusted. This is iterated until convergence.

1.7 HOST-GUEST INTERACTIONS

While these methods exist to guide force-field parameterization, a major limitation still exists in the form of choosing data for the optimization process. As mentioned previously, successful efforts in the past have used neat liquids to generate experimental data to tune parameters[24], but this possibly limits the applications, as the parameters

generated may not transfer well to a binding calculation, for example. Thus, in order to improve force-fields, particularly for use in estimating binding free energies, it may be imperative to include binding data into the training set of data used to optimize parameters. However, protein-ligand binding models are exceedingly complicated, and often require great deals of time to simulate enough data to ensure accurate results. A simpler model is required in order to maximize speed, but still requires a non-covalent binding interaction.

A common solution to this issue is use of host-guest interactions[71], which typically refer to supramolecular interactions involving small macromolecules such as the glucose polymer cyclodextrin[72], a ring of six to eight glucose molecules. The cyclodextrin family of molecules have a hydrophobic cavity in the center which binds small molecules from water, effectively reproducing the non-covalent binding behavior of a small molecule binding a pocket on a protein. While this behavior is preserved, the system itself is much smaller and simpler, allowing both the simulation time and setup to be more straightforward. This allows the researcher to ensure that most inaccuracies are coming from the force-field and associated parameters, and not from errors associated with sampling or the size of the system being studied. Indeed, host-guest complexes have been used as a benchmark to examine accuracy in binding calculation accuracy in multiple iterations of the SAMPL challenge[27,31,58].

It is often straightforward to measure the binding free energies and enthalpies of these systems, and thus provide data to tune force field parameters. However, the standard cyclodextrin hosts offer little chemical diversity, relative to what would be most desirable for use in testing and adjusting force fields. . Thus, a method to modify host-guest pairs with modular functional groups would be of great use to expanding data sets

used to optimize force-field parameters. In Chapter 3, I will discuss some of my work aimed at functionalizing host molecules such as beta-cyclodextrin with various groups that can expand the diversity of a force field training set.

1.8 ISOTHERMAL TITRATION CALORIMETRY

Once a set of host-guest pairs has been procured or synthesized, the challenge of accurately measuring the binding thermodynamics remains. One of the most reliable and informative choice of experimental measurements is isothermal titration calorimetry (ITC)[73], which provides a direct, simultaneous measurement of both the binding free energy and binding enthalpy, thus providing a rich set of information about the thermodynamics of binding. In contrast, most other methods, such as ones based on spectroscopy, provide only the binding free energy directly; the enthalpy must be derived by the error-prone technique of repeated affinity measurements at multiple temperatures, followed by application of the van 't Hoff equation[74], as further discussed below.

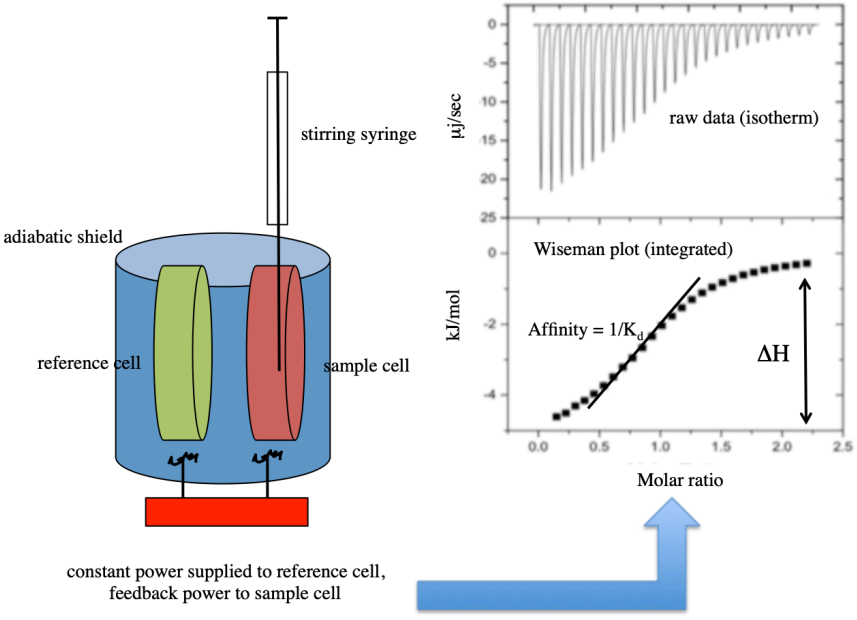


Figure 1.3: Overview of isothermal titration calorimetry; the left hand side shows a simple schematic of the calorimeter, demonstrating how titrant is injected into the cell. The right hand side shows the result of a series of injections, and the resultant integrated peaks (Wiseman plot).

ITC measurements work by injecting small aliquots of solution containing one molecule into a cell containing a solution of another molecule. If the two molecules bind, the heat either taken up or released into the cell solution during the binding reaction and is measured. The instrument uses power to keep the sample cell at the same temperature as a matched reference cell, and the change in that power causes peaks in the thermogram (Figure 1.3). These peaks are integrated to get the heat released for each injection, providing a binding isotherm called a Wiseman plot. When combined with experimental details like the concentrations of the cell and syringe reactants and the injection volumes, theseplots can then be fit to provide both the binding enthalpy and the affinity or binding free energy.

One serious consideration when using ITC data is the propagation of error, and the reliability of the provided data[75]. Because thermodynamics obtained from simulations are so sensitive to force-field parameters, even small errors in the data used to train parameters may lead to serious errors. Thus, a complete understanding of how to appropriately fit the data and how the sources of experimental noise affect the goodness of the fit to the data are required if the data is to be used for parameterization. Chapter 4 provides a novel and detailed analysis of the sources, character, and magnitude of the uncertainties in ITC measurements of binding free energy and enthalpy.

There has also been considerable controversy regarding the reported mismatches between calorimetric and van't Hoff enthalpies[76–79]. A calorimetric enthalpy is calculated directly from the ITC data, whereas a van't Hoff enthalpy is calculated from the change in binding free energy over some temperature range. In theory, the two should match, and mismatches could possibly inform errors in the methodology. This relationship is of particular interest to us as we sought to look at how experimental errors

affect data, and the mismatch between the two enthalpies might provide information as to the level and sources of error. This topic is considered in Chapter 5.

1.9 TOWARDS IMPROVED FORCE-FIELDS FOR BINDING CALCULATIONS

Altogether, this thesis aims to provide building blocks for a means to systematically improve force-fields for binding calculations. The first step to consider is the optimization of the force-field. To this end, we developed the Slater-Derived Lennard Jones (SDLJ) method, which is detailed in Chapter 2. This method utilizes information from quantum mechanical calculations to generate physically motivated LJ terms with only a few adjustable parameters. The next consideration is data to be used in force-field optimization. As discussed above, the intent is to incorporate experimental binding data into the optimization process, so as to make a given set of parameters more relevant for binding calculations. To accomplish this, we set out to design a modifiable host macromolecule in a systematic and controlled way, so as to maximize control over the design of the structure. Accordingly, Chapter 3 details the synthesis of a monopropargylated beta-cyclodextrin, which serves as a starting point for derivatization by click chemistry. Finally, we require a reliably accurate method to measure the experimental binding data. Isothermal titration calorimetry was chosen for the advantages given above; however we required a more detailed understanding of the error associated with the instrument and the process of analyzing the data. Additionally, we sought to resolve the issues between van't Hoff and calorimetric-derived enthalpies of binding, as the discrepancies observed had not been suitably explained, and possibly yield information about the error. These studies are detailed in Chapters 4 and 5.

1.10 REFERENCES

1. Griebel, M., Knapek, S. & Zumbusch, G. *Numerical Simulation in Molecular Dynamics: Numerics, Algorithms, Parallelization, Applications*. (Springer Publishing Company, Incorporated, 2007).

2. Allen, M. P. & Tildesley, D. J. Computer Simulation of Liquids. in (1988).

3. Froese Fischer, C. General Hartree-Fock program. *Comput. Phys. Commun.* **43**, 355–365 (1987).

4. Szabo, A. & Ostlund, N. S. *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*. (Courier Corporation, 2012).

5. Dewar, M. J. S., Zoebisch, E. G., Healy, E. F. & Stewart, J. J. P. Development and use of quantum mechanical molecular models. 76. AM1: a new general purpose quantum mechanical molecular model. *J. Am. Chem. Soc.* **107**, 3902–3909 (1985).

6. Brooks, B. R., Bruccoleri, R., Olafson, B., States, D., Swaminathan, S. CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem.* **4**, 187–217 (1983).

7. Van Der Spoel, D., Lindahl E., Hess B., Groenhof G., Mark A.E., Berendsen H.J. GROMACS: Fast, flexible, and free. *J. Comput. Chem.* **26**, 1701–1718 (2005).

8. Salomon-Ferrer, R., Case, D. A. & Walker, R. C. An overview of the Amber biomolecular simulation package. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **3**, 198–210 (2013).

9. Kosztolányi, T., Bakó, I. & Palinkas, G. Hydrogen bonding in liquid methanol, methylamine, and methanethiol studied by molecular-dynamics simulations. *J. Chem. Phys.* **118**, 4546–4555 (2003).

10. Saiz, L., Padró, J. A. & Guàrdia, E. Structure and Dynamics of Liquid Ethanol. *J. Phys. Chem. B* **101**, 78–86 (1997).

11. Ryckaert, J.-P. & Bellemans, A. Molecular dynamics of liquid alkanes. *Faraday Discuss. Chem. Soc.* **66**, 95–106 (1978).

12. Durrant, J. D. & McCammon, J. A. Molecular dynamics simulations and drug discovery. *BMC Biol.* **9**, 71 (2011).

13. Homeyer, N., Stoll, F., Hillisch, A. & Gohlke, H. Binding Free Energy Calculations for Lead Optimization: Assessment of Their Accuracy in an Industrial Drug Design Context. *J. Chem. Theory Comput.* **10**, 3331–3344 (2014).

14. Klepeis, J. L., Lindorff-Larsen, K., Dror, R. O. & Shaw, D. E. Long-timescale molecular dynamics simulations of protein structure and function. *Curr. Opin. Struct. Biol.* **19**, 120–127 (2009).

15. Raval, A., Piana, S., Eastwood, M. P., Dror, R. O. & Shaw, D. E. Refinement of protein structure homology models via long, all-atom molecular dynamics simulations. *Proteins Struct. Funct. Bioinforma.* n/a-n/a (2012).

16. Yamakov, V., Wolf, D., Phillpot, S. R., Mukherjee, A. K. & Gleiter, H. Dislocation processes in the deformation of nanocrystalline aluminium by molecular-dynamics simulation. *Nat. Mater.* **1**, 45–49 (2002).

17. Yamakov, V., Wolf, D., Phillpot, S. R., Mukherjee, A. K. & Gleiter, H. Deformation-mechanism map for nanocrystalline metals by molecular-dynamics simulation. *Nat. Mater.* **3**, 43–47 (2004).

18. Lustig, R. Statistical thermodynamics in the classical molecular dynamics ensemble. I. Fundamentals. *J. Chem. Phys.* **100**, 3048–3059 (1994).

19. Wang, L.-P., Martinez, T. J. & Pande, V. S. Building Force Fields: An Automatic, Systematic, and Reproducible Approach. *J. Phys. Chem. Lett.* **5**, 1885–1891 (2014).

20. Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A. & Case, D. A. Development and testing of a general amber force field. *J. Comput. Chem.* **25**, 1157–1174 (2004).

21. Faller, R., Schmitz, H., Biermann, O. & Müller-Plathe, F. Automatic parameterization of force fields for liquids by simplex optimization. *J. Comput. Chem.* **20**, 1009–1017 (1999).

22. Betz, R. M. & Walker, R. C. Paramfit: automated optimization of force field parameters for molecular dynamics simulations. *J. Comput. Chem.* **36**, 79–87 (2015).

23. Horton, J. T., Allen, A. E. A., Dodda, L. S. & Cole, D. J. QUBEKit: Automating the Derivation of Force Field Parameters from Quantum Mechanics. *J. Chem. Inf. Model.* **59**, 1366–1381 (2019).

24. Jorgensen, W. L., Maxwell, D. S. & Tirado-Rives, J. Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. *J. Am. Chem. Soc.* **118**, 11225–11236 (1996).

25. Ponder, J. W., Wu C., Ren P., Pande V.S., Chodera J.D., Schnieders M.J., Haque I., Mobley D.L., Lambrecht D.S., DiStasio R.A. Jr., Head-Gordon M., Clark G.N., Johnson M.E., Head-Gordon T. Current Status of the AMOEBA Polarizable Force Field. *J. Phys. Chem. B* **114**, 2549–2564 (2010).

26. Laury, M. L., Wang, L.-P., Pande, V. S., Head-Gordon, T. & Ponder, J. W. Revised Parameters for the AMOEBA Polarizable Atomic Multipole Water Model. *J. Phys. Chem. B* **119**, 9423–9437 (2015).

27. Skillman, A. G. SAMPL3: blinded prediction of host–guest binding affinities, hydration free energies, and trypsin inhibitors. *J. Comput. Aided Mol. Des.* **26**, 473–474 (2012).

28. Yin, J., Henriksen, N. M., Slochower, D. R. & Gilson, M. K. The SAMPL5 host–guest challenge: computing binding free energies and enthalpies from explicit solvent simulations by the attach-pull-release (APR) method. *J. Comput. Aided Mol. Des.* **31**, 133–145 (2017).

29. Bosisio, S., Mey, A. S. J. S. & Michel, J. Blinded predictions of host-guest standard free energies of binding in the SAMPL5 challenge. *J. Comput. Aided Mol. Des.* **31**, 61–70 (2017).

30. Geballe, M. T., Skillman, A. G., Nicholls, A., Guthrie, J. P. & Taylor, P. J. The SAMPL2 blind prediction challenge: introduction and overview. *J. Comput. Aided Mol. Des.* **24**, 259–279 (2010).

31. Muddana, H. S., Varnado C.D., Bielawski C.W., Urbach A.R., Isaacs L, Geballe M.T., Gilson M.K. Blind prediction of host-guest binding affinities: a new SAMPL3 challenge. *J. Comput. Aided Mol. Des.* **26**, 475–487 (2012).

32. Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., DiNola, A. & Haak, J. R. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **81**, 3684–3690 (1984).

33. Posch, H. A., Hoover, W. G. & Vesely, F. J. Canonical dynamics of the Nosé oscillator: Stability, order, and chaos. *Phys. Rev. A* **33**, 4253–4265 (1986).

34. Schlick, T. *Molecular Modeling and Simulation: An Interdisciplinary Guide: An Interdisciplinary Guide*. (Springer Science & Business Media, 2010).

35. Kolafa, J. & Perram, J. W. Cutoff Errors in the Ewald Summation Formulae for Point Charge Systems. *Mol. Simul.* **9**, 351–368 (1992).

36. Margenau, H. & Kestner, N. R. *Theory of intermolecular forces*. (Pergamon Press, 1971).

37. Eisenschitz, R. & London, F. Über das Verhältnis der van der Waalsschen Kräfte zu den homöopolaren Bindungskräften. *Z. Für Phys.* **60**, 491–527 (1930).

38. Snow, C. D., Zagrovic, B. & Pande, V. S. The Trp Cage: Folding Kinetics and Unfolded State Topology via Molecular Dynamics Simulations. *J. Am. Chem. Soc.* **124**, 14548–14549 (2002).

39. Swope, W. C., Pitera, J. W. & Suits, F. Describing Protein Folding Kinetics by Molecular Dynamics Simulations. 1. Theory †. *J. Phys. Chem. B* **108**, 6571–6581 (2004).

40. Marrink, S. J., Tieleman, D. P. & Mark, A. E. Molecular Dynamics Simulation of the Kinetics of Spontaneous Micelle Formation. *J. Phys. Chem. B* **104**, 12165–12173 (2000).

41. Kaminski, G., Duffy, E. M., Matsui, T. & Jorgensen, W. L. Free Energies of Hydration and Pure Liquid Properties of Hydrocarbons from the OPLS All-Atom Model. *J. Phys. Chem.* **98**, 13077–13082 (1994).

42. Gilson, M. K., Given, J. A., Bush, B. L. & McCammon, J. A. The statistical-thermodynamic basis for computation of binding affinities: a critical review. *Biophys. J.* **72**, 1047–1069 (1997).

43. Henriksen, N. M., Fenley, A. T. & Gilson, M. K. Computational Calorimetry: High-Precision Calculation of Host–Guest Binding Thermodynamics. *J. Chem. Theory Comput.* **11**, 4377–4394 (2015).

44. Frenkel, D. & Smit, B. *Understanding molecular simulation: from algorithms to applications*. (Academic Press, 2002).

45. Zwanzig, R. W. High-Temperature Equation of State by a Perturbation Method. I. Nonpolar Gases. *J. Chem. Phys.* **22**, 1420–1426 (1954).

46. Bennett, C. H. Efficient estimation of free energy differences from Monte Carlo data. *J. Comput. Phys.* **22**, 245–268 (1976).

47. Shirts, M. R. & Chodera, J. D. Statistically optimal analysis of samples from multiple equilibrium states. *J. Chem. Phys.* **129**, (2008).

48. Barducci, A., Bonomi, M. & Parrinello, M. Metadynamics. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **1**, 826–843 (2011).

49. Bussi, G., Gervasio, F. L., Laio, A. & Parrinello, M. Free-energy landscape for beta hairpin folding from combined parallel tempering and metadynamics. *J. Am. Chem. Soc.* **128**, 13435–13441 (2006).

50. Limongelli, V., Bonomi, M. & Parrinello, M. Funnel metadynamics as accurate binding free-energy method. *Proc. Natl. Acad. Sci.* **110**, 6358–6363 (2013).

51. Hamelberg, D., Mongan, J. & McCammon, J. A. Accelerated molecular dynamics: A promising and efficient simulation method for biomolecules. *J. Chem. Phys.* **120**, 11919–11929 (2004).

52. Miao, Y., Feher, V. A. & McCammon, J. A. Gaussian Accelerated Molecular Dynamics: Unconstrained Enhanced Sampling and Free Energy Calculation.

53. Yang, Y. I., Shao, Q., Zhang, J., Yang, L. & Gao, Y. Q. Enhanced sampling in molecular dynamics. *J. Chem. Phys.* **151**, 070902 (2019).

54. Shaw, D. E. Millisecond-scale molecular dynamics simulations on Anton. 11.

55. Shaw, D. E. Anton 2: Raising the Bar for Performance and Programmability in a Special-Purpose Molecular Dynamics Supercomputer. in *SC14: International Conference for High Performance Computing, Networking, Storage and Analysis* 41–53 (2014).

56. Lee, T. Cerutti D.S., Mermelstein D., Lin C., LeGrand S., Giese T.J., Roitberg A., Case D.A., Walker R.C., York D.M. GPU-accelerated molecular dynamics and free energy methods in Amber18: performance enhancements and new features. *J. Chem. Inf. Model.* (2018)

57. Salomon-Ferrer, R., Götz, A. W., Poole, D., Le Grand, S. & Walker, R. C. Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 2. Explicit Solvent Particle Mesh Ewald. *J. Chem. Theory Comput.* **9**, 3878–3888 (2013).

58. Muddana, H. S. & Gilson, M. K. Prediction of SAMPL3 host–guest binding affinities: evaluating the accuracy of generalized force-fields. *J. Comput. Aided Mol. Des.* **26**, 517–525 (2012).

59. Caleman, C. van Maaren P.J., Hong M., Hub J.S., Costa L.T., van der Spoel D. Force Field Benchmark of Organic Liquids: Density, Enthalpy of Vaporization, Heat Capacities, Surface Tension, Isothermal Compressibility, Volumetric Expansion Coefficient, and Dielectric Constant. *J. Chem. Theory Comput.* **8**, 61–74 (2012).

60. Freitag, M. A., Gordon, M. S., Jensen, J. H. & Stevens, W. J. Evaluation of charge penetration between distributed multipolar expansions. *J. Chem. Phys.* **112**, 7300–7306 (2000).

61. Wang, B. & Truhlar, D. G. Including Charge Penetration Effects in Molecular Modeling. *J. Chem. Theory Comput.* **6**, 3330–3342 (2010).

62. Lim, J. Y. C. & Beer, P. D. Sigma-Hole Interactions in Anion Recognition. *Chem* **4**, 731–783 (2018).

63. Quiñonero, D. Sigma-hole carbon-bonding interactions in carbon–carbon double bonds: an unnoticed contact. *Phys. Chem. Chem. Phys.* **19**, 15530–15540 (2017).

64. Van Vleet, M. J., Misquitta, A. J., Stone, A. J. & Schmidt, J. R. Beyond Born-Mayer: Improved Models for Short-Range Repulsion in ab Initio Force Fields. *J. Chem. Theory Comput.* **12**, 3851–3870 (2016).

65. Mobley, D. L. Bannan C.C., Rizzi A., Bayly C.I., Chodera J.D., Lim V.T., Lim N.M., Beauchamp K.A., Slochower D.R., Shirts M.R., Gilson M.K., Eastman P.K. Escaping Atom Types in Force Fields Using Direct Chemical Perception. *J. Chem. Theory Comput.* **14**, 6076–6092 (2018).

66. Tkatchenko, A. & Scheffler, M. Accurate Molecular Van Der Waals Interactions from Ground-State Electron Density and Free-Atom Reference Data. *Phys. Rev. Lett.* **102**, 073005 (2009).

67. Cole, D. J., Vilseck, J. Z., Tirado-Rives, J., Payne, M. C. & Jorgensen, W. L. Biomolecular Force Field Parameterization via Atoms-in-Molecule Electron Density Partitioning. *J. Chem. Theory Comput.* **12**, 2312–2323 (2016).

68. Li, Y. Li H., Pickard F.C., Narayanan B., Sen F.G., Chan M.K.Y., Sankaranarayanan S.K.R.S., Brooks B.R., Roux B. Machine Learning Force Field Parameters from Ab Initio Data. *J. Chem. Theory Comput.* **13**, 4492–4503 (2017).

69. Yin, J., Fenley, A. T., Henriksen, N. M. & Gilson, M. K. Toward Improved Force-Field Accuracy through Sensitivity Analysis of Host-Guest Binding Thermodynamics. *J. Phys. Chem. B* **119**, 10145–10155 (2015).

70. Verstraelen, T., Vandenbrande S., Heidar-Zadeh F., Vanduyfhuys L., Van Speybroeck V., Waroquier M., Ayers P.W. Minimal Basis Iterative Stockholder: Atoms in Molecules for Force-Field Development. *J. Chem. Theory Comput.* **12**, 3894–3912 (2016).

71. Cram, D. J. & Cram, J. M. Host-Guest Chemistry: Complexes between organic compounds simulate the substrate selectivity of enzymes. *Science* **183**, 803–809 (1974).

72. Rekharsky, M. V. & Inoue, Y. Complexation Thermodynamics of Cyclodextrins. *Chem. Rev.* **98**, 1875–1918 (1998).

73. Wiseman, T., Williston, S., Brandts, J. F. & Lin, L.-N. Rapid measurement of binding constants and heats of binding using a new titration calorimeter. *Anal. Biochem.* **179**, 131–137 (1989).

74. Hoff, J. H. van't (Jacobus H. van't). *Etudes de dynamique chimique*. (Amsterdam : Frederik Muller, 1884).

75. Tellinghuisen, J. & Chodera, J. D. Systematic errors in isothermal titration calorimetry: concentrations and baselines. *Anal. Biochem.* **414**, 297–299 (2011).

76. Chaires, J. B. Possible origin of differences between van't Hoff and calorimetric enthalpy estimates. *Biophys. Chem.* **64**, 15–23 (1997).

77. Horn, J. R., Russell, D., Lewis, E. A. & Murphy, K. P. Van't Hoff and calorimetric enthalpies from isothermal titration calorimetry: are there significant discrepancies? *Biochemistry* **40**, 1774–1778 (2001).

78. Horn, J. R., Brandts, J. F. & Murphy, K. P. van't Hoff and calorimetric enthalpies II: effects of linked equilibria. *Biochemistry* **41**, 7501–7507 (2002).

79. Castellano, B. M. & Eggers, D. K. Experimental Support for a Desolvation Energy Term in Governing Equations for Binding Equilibria. *J. Phys. Chem. B* **117**, 8180–8188 (2013).

2 DATA-DRIVEN MAPPING OF GAS-PHASE QUANTUM CALCULATIONS TO GENERAL FORCE FIELD LENNARD-JONES PARAMETERS

JCTC Journal of Chemical Theory and Computation

pubs.acs.org/JCTC Article

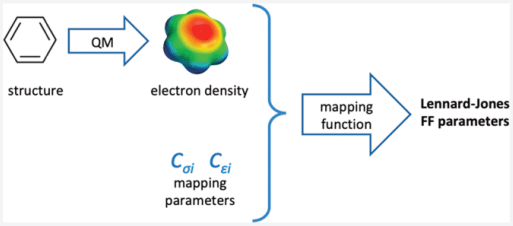
Data-Driven Mapping of Gas-Phase Quantum Calculations to General Force Field Lennard-Jones Parameters

Sophie M. Kantonen, Hari S. Muddana, Michael Schauperl, Niel M. Henriksen, Lee-Ping Wang, and Michael K. Gilson*

Cite This: https://dx.doi.org/10.1021/acs.jctc.9b00713 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Molecular dynamics simulations are helpful tools for a range of applications, ranging from drug discovery to protein structure determination. The successful use of this technology largely depends on the potential function, or force field, used to determine the potential energy at each configuration of the system. Most force fields encode all of the relevant parameters to be used in distinct atom types, each associated with parameters for all parts of the force field, typically bond stretches, angle bends, torsions, and nonbonded terms accounting for van der Waals and electrostatic interactions. Much attention has been paid to the nonbonded parameters and their derivation, which are important in particular due to their governance of noncovalent interactions, such as protein-ligand binding. Parametrization involves adjusting the nonbonded parameters to minimize the error between simulation results and experimental properties, such as heats of vaporization and densities of neat liquids. In this setting, determining the best set of atom types is far from trivial, and the large number of parameters to be fit for the atom types in a typical force field can make it difficult to approach a true optimum. Here, we utilize a previously described Minimal Basis Iterative Stockholder (MBIS) method to carry out an atoms-in-molecules partitioning of electron densities. Information from these atomic densities is then mapped to Lennard-Jones parameters using a set of mapping parameters much smaller than the typical number of atom types in a force field. This approach is advantageous in two ways: it eliminates atom types by allowing each atom to have unique Lennard-Jones parameters, and it greatly reduces the number of parameters to be optimized. We show that this approach yields results comparable to those obtained with the typed GAFF 1.7 force field, even when trained on a relatively small amount of experimental data.



1. INTRODUCTION

Molecular dynamics (MD) simulations are useful for a broad range of applications, including protein folding studies,[1] drug discovery,[2] and the determination of liquid structure and properties.[3] Various approaches have been taken to improve the ability of simulations to explore the thermodynamically relevant parts of configuration space, including hardware advancements[4−9] and more effective sampling algorithms.[10−14] While these efforts have dramatically improved our ability to generate well-converged results, errors persist in simulations, as highlighted for example, in the SAMPL series of blinded prediction exercises.[15−21] Therefore, attention is now turning again to the potential functions, or force fields (FF), as sources of error, as recently reviewed.[22]

Currently, most biomolecular simulations are still carried out with FFs having a simple functional form comprising harmonic bond-stretches and angle-bends, sinusoidal torsional terms, harmonic improper dihedrals, Lennard-Jones (LJ) interactions for van der Waals forces, and Coulombic interactions among atom-centered point charges for electrostatic interactions and hydrogen bonding.[23] This common functional form has the merit of being supported by many well-developed simulation packages and of affording great computational speed and therefore effective conformational sampling. From this starting point, a number of strategies may be adopted to improve the accuracy of the FF. One is to move to a functional form that captures the physics more faithfully and in more detail. For example, one may add terms that explicitly account for effects, such as electronic polarizability,[24−27] that are accounted for only implicitly, at best, in the common functional form; or one may substitute a more realistic form for an existing term, such as a Coulombic term that accounts for charge penetration.[26−31] Another strategy is to remain with today's common functional form and instead look for parameters, such as atomic partial charges, torsional barriers, and Lennard-Jones well-depths and radii, that will lead to greater accuracy when used to compute

Received: July 16, 2019
Published: January 9, 2020

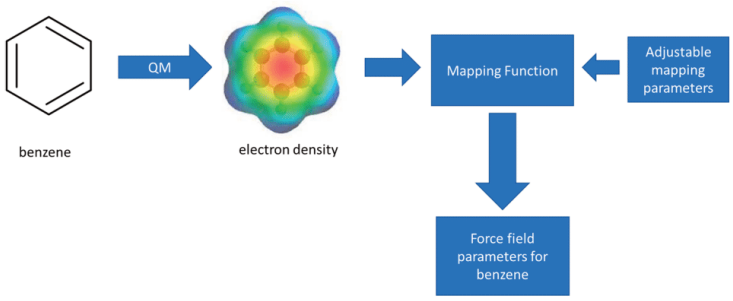


Figure 1. Generic scheme for use of a parametrized mapping from QM results to FF parameters, with benzene as an example.

experimental and/or quantum chemical reference data. This strategy may be pursued by using more, and more relevant, experimental data in the parametrization process. For example, data on host−guest binding thermodynamics have recently been used to guide parameter adjustment, leading to improved accuracy in calculated binding thermodynamics. Another approach to improving accuracy of force fields with the common functional form is to more comprehensively and systematically define[32] and adjust[33−38] the parameters so they more closely approach an optimal parametrization against a fixed data set.

However, the large number of independently adjustable FF parameters in a typical FF can make full, multidimensional parameter optimization daunting. For example, the SMIRNOFF99Frost[39] FF, whose list of parameters has already been condensed through the replacement of atom-typing with direct chemical perception, still has 35 different Lennard-Jones types and hence 70 Lennard-Jones parameters. Simultaneous optimization of these parameters becomes particularly challenging when evaluating the objective function requires running time-consuming simulations, such as if one wishes to tune parameters against liquid-state properties. Even more extensive sampling of parameters will likely be needed if one moves from optimization to Bayesian sampling[40−42] in the parameter space. If either an optimization or a Bayes sampling algorithm misses a key sector of parameter space, the accuracy of the resulting parametrized FF will not be a measure of the quality achievable within the common functional form. This situation is problematic for at least two reasons. First, simulations using the incompletely optimized FF will not be as accurate as they could have been. Second, the resulting errors may provide misleading guidance regarding the need to move to a more complex functional form. Consequently, a methodology of fitting FFs that use fewer adjustable parameters would in principle make the problem more tractable.

We therefore propose a step toward reducing the dimensionality of the adjustable parameter space of the common FF functional form. The basic idea (Figure 1) is to use a QM calculation on the molecule to be parametrized (or on a suitable fragment of a large molecule) to extract properties of the electron density that correlate with the FF parameters to be assigned—here the Lennard-Jones parameters of each atom in the molecule. We then set up a mathematical mapping from the electron density to the targeted FF parameters. The mapping is outfitted with a small set of adjustable parameters, and it is these mapping parameters, not the targeted FF parameters, that are subjected to optimization or sampling based on calculations of experimental observables. This approach reduces the number of adjustable parameters, because the FF parameters assigned to each atom in the molecule are largely controlled by the QM results. Only the mapping parameters are adjusted to maximize the agreement of simulated properties with experiment. Thus, the dimensionality of the optimization problem is greatly reduced. By the same token, this approach allows each atom in a molecule to have unique FF parameters without requiring a large number of atom types.

The present study aims to prove the principle of this approach by demonstrating that such a QM-to-FF mapping, trained to generate Lennard-Jones FF parameters that best replicate a small set of experimental liquid-state data, yields a competitive level of accuracy for an informative set of properties. It thus sets the stage for future applications aimed at building a comprehensive FF for general use. The current implementation builds on promising approaches from other groups,[28,43−50] as it uses Slater orbitals to model the electron density associated with each atom in the molecule and extracts key correlates of the LJ parameters from these fits. The present FF-development approach is therefore termed the Slater-Derived Lennard-Jones (SDLJ) method. The following subsections detail the concepts and methodology, describe how the mapping parameters are adjusted against experimental liquid-state data for a small training set of compounds, and report on the quality of the results when the trained method is tested on a larger, nonoverlapping set of compounds and additional observables. Implications and prospects are considered in the Discussion section.

2. METHODS

2.1. Overview. The present method provides an approach to mapping from the electronic structure of a molecule, obtained from a quantum mechanical (QM) calculation, to suitable σ and ε Lennard-Jones parameters for each of its atoms. The mapping contains two adjustable parameters for each element (C, N, O, H), except that polar and nonpolar hydrogens have separate parameters. Here a polar hydrogen is defined simply as one directly bound to a nitrogen or oxygen atom. We separated polar and nonpolar hydrogens because initial studies showed that lumping all hydrogens made it impossible to reach a set of LJ parameters that would afford competitive accuracy (results not shown). We use the ForceBalance software[34] to optimize the 10 mapping parameters in order to minimize the error of simulated liquid

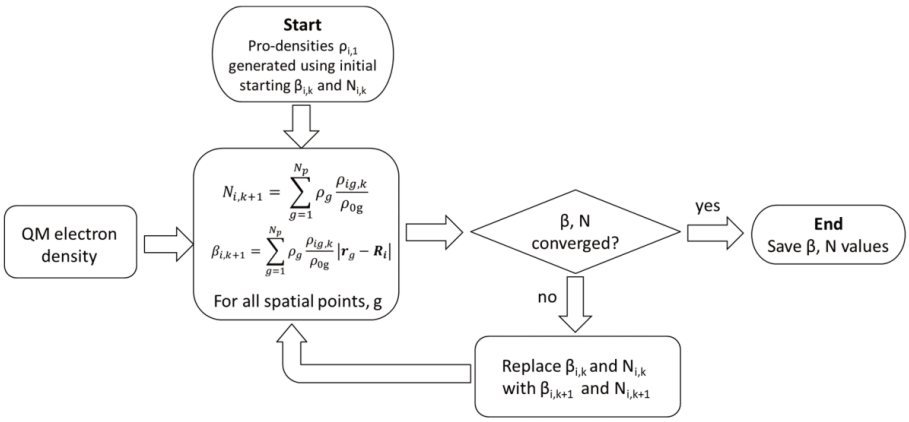


Figure 2. Schematic of the iterative stockholder method used to fit β and N for each atom-in-molecule i to the molecular electron density from a QM calculation. Symbols are defined in the main text.

state properties vs experimental measurements. The trained mapping is then tested against the properties of a larger set of liquids, and the results are compared with those obtained with the pioneering and widely used GAFF force field.[51] For this proof-of-concept study, we focused on consequences of adjusting only Lennard-Jones parameters, thus the electrostatic interactions are modeled with atom-center partial charges assigned with the AM1/BCC method and all non-LJ force field terms were drawn from GAFF 1.7 in accordance with established procedures.

The mapping from a molecule's electronic structure to σ_i and ε_i for each of its atoms, i, uses an atoms-in-molecules (AIM) approach similar, but not identical, to ones that have been published before.[43,44] We model the electron density around each atom i in terms of a Slater orbital, where the electron density decays exponentially with distance from the nucleus: $\rho_i(r) \propto \exp(-\beta_i r)$. The decay coefficient β_i associated with each atom i then is used to assign both the effective atomic polarizability and the effective ionization potential, both key quantities in the LJ interaction term. The following subsections detail each step summarized in this overview. The code used to go from a QM result to Lennard-Jones parameters is available at (https://github.com/SKantonen/PyBLJ). It is written in Python, and its inputs are a Lebedev grid file and a mol2 file to generate β_i coefficients for each atom in the given mol2 structure. The level of QM to be run can be manually set inside of the code, if so desired.

2.2. Electron Density Calculations. Each molecule to be studied was built with the open-source software Avogadro,[52] and its structure was energy-minimized using the GAFF 1.7 force field. The Gaussian09[53] package was used to obtain the electron density at the CCSD/cc-pVTZ level of theory. CCSD was chosen as it is considered to be among the most accurate post Hartree−Fock methods for calculations on small molecules.[54] We found that the values of β changed by <1% when the size of the basis set was increased further. The electron density distribution around each atom was computed using spherical Lebedev grids (110 points, order 17)[55] and a uniform radial grid of 0.05 Å spacing out from 0−12 Å. The code used to generate these grids is built into the β parameter fitting code, using the aforementioned spacings and grid sizes.

2.3. Fitting Electron-Density Decay Coefficients (Beta) to Atoms-in-Molecules. While there exist multiple useful electron partitioning methods,[56−58] we chose to use the Minimal Basis Iterative Stockholder (MBIS) method of Verstraelen.[59] The MBIS method partitions the total electron density of the molecule (Section 2.2) into atom-centered Slater orbitals centered on atoms i, in a manner that minimizes the Kullback-Liebler (KL) divergence[59] between the electron density distribution provided by a QM calculation $\rho(\mathbf{r})$ and the sum of the atomic densities. The MBIS method is attractive because it allows pro-densities to vary (allowing individual atoms to have unique parameters governing their pro-densities as opposed to predetermined pro-densities) and has already been successfully applied to force field development.[46,59]

Thus, each atom, i, is assigned a "pro-density", $\rho_i(\mathbf{r})$, of Slater form whose integral over all space gives the total number of electrons assigned to the atom, N_i. The Slater orbital of atom i is characterized by β_i, the spatial decay constant of the electron density.[60] This quantity is expected, on physical grounds, to correlate with both the size and the dispersion interactions of the atom,[60,61] as detailed below.

As previously shown, the KL divergence is minimized by an iterative procedure, diagrammed in Figure 2. For a given iteration, k, of the MBIS method

$$N_{i,k+1} = \int \rho(\mathbf{r}) \frac{\rho_{i,k}(\mathbf{r}, \beta_{i,k}, N_{i,k})}{\rho_{0,k}(\mathbf{r})} \, d\mathbf{r} \tag{1}$$

$$\beta_{i,k+1} = 3N_{i,k} \int \rho(\mathbf{r}) \frac{\rho_{i,k}(\mathbf{r}, \beta_{i,k}, N_{i,k})}{\rho_{0,k}(\mathbf{r})} |\mathbf{r} - \mathbf{R}_i| \, d\mathbf{r} \tag{2}$$

Here $\rho_{i,k}(\mathbf{r})$ is the pro-density around atom i at iteration k; $\beta_{i,k}$ is the estimate of β_i for atom i at iteration k; $\rho_0(\mathbf{r})$ is the sum of all atomic pro-densities at position $\mathbf{r}$; and $\mathbf{R}_i$ is the location of the nucleus of atom i. In practice, these integrals are estimated as sums over Lebedev grids, as noted above:

$$N_{i,k+1} = \sum_{g=1}^{N_p} \rho(\mathbf{r}_g) \frac{\rho_{i,k}(\mathbf{r}_g, \beta_{i,k}, N_{i,k})}{\rho_0(\mathbf{r}_g)} \tag{3}$$

$$\beta_{i,k+1} = \sum_{g=1}^{N_p} \rho(\mathbf{r}_g) \frac{\rho_{i,k}(\mathbf{r}_g, \beta_{i,k}, N_{i,k})}{\rho_g(\mathbf{r}_g)} |\mathbf{r}_g - \mathbf{R}_i| \tag{4}$$

For the sums, the density at each grid point g is considered, with the sum being over all N_p grid points, each located at $\mathbf{r}_g$. As per the MBIS method, an initial value, corresponding to iteration $k = 1$, is chosen for all N_{i1} and β_{i1}, allowing for the determination of starting pro-densities via

$$\rho_{i1}(\mathbf{r}) = \frac{N_{i1}\beta_i^3}{8\pi} e^{(-\beta_{i1}|\mathbf{r}-\mathbf{R}_i|)} \tag{5}$$

These pro-densities are used to generate $N_{i,k}$ and $\beta_{i,k}$ for the next iteration ($k = 2$), and the process is iterated until the changes in N and β between iterations k and $k+1$ fall below some threshold, here a 0.05% absolute change. The initial values of these quantities are detailed in the following paragraph.

In accord with Verstraelen et al., we assigned and fit two pro-densities to each non-hydrogen atom. (Only one pro-density is used for each hydrogen atom.) The "core" pro-density captures electrons held close to the nucleus, while the "valence" one includes all other electrons, and the number of electrons in core versus valence is adjusted as part of the procedure. Thus, we allow two β's and two N's to be fit for each non-hydrogen atom, and we use only the valence β to generate LJ parameters. The rationale is that core electrons only contribute to nonbonded interactions at distances much too close to be relevant, mostly due to exchange repulsion.[61] The initial values of N are set to 2 for the core pro-density and the number of valence electrons of the element for the valence pro-density. Initial values of β_i for all atoms are set to 12 and 4 Å^{-1} for the core and valence orbitals, respectively. These values, which correspond to those obtained for a single nitrogen atom, suffice to generate convergent results in the iterative procedure just outlined. As expected, core occupancy was on average very near two electrons, but the addition of the core exponential allowed for a better fit of the valence exponential. Values of β typically converge relatively quickly, with the final value being insensitive to the starting value, as shown in Supplementary Figure 1.

2.4. Mapping QM Results to Lennard-Jones Parameters. The Lennard-Jones model gives the van der Waals interaction energy between atoms i and j as

$$E_{LJ} = \frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^{6}} = 4\varepsilon_{ij}\left(\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right) \tag{6}$$

where A_{ij}, B_{ij}, σ_{ij}, and ε_{ij} may be obtained from atomic "self" parameters σ_i, σ_j, ε_i, and ε_j by mixing rules, such as $\sigma_{ij} = 0.5(\sigma_i + \sigma_j)$ and $\varepsilon_{ij} = (\varepsilon_i\varepsilon_j)^{1/2}$.[62] The next two subsections describe how σ_i and ϵ_i are assigned to each atom in a molecule, based on the QM electron densities.

2.4.1. Lennard-Jones Sigma. The parameter σ is essentially an atomic diameter, and β^{-1} is proportional to the expectation value of the distance of the electron density from the nucleus,[59] so we write

$$\sigma_i = \frac{C_{\sigma,\epsilon_i}}{\beta_i} \tag{7}$$

Here ϵ_i is the element of atom i, and C_{σ,ϵ_i} is the associated element-specific mapping parameter, which is adjusted with ForceBalance[34] (Section 2.5).

2.4.2. Lennard-Jones Epsilon. The coefficient of the dispersion component of the Lennard-Jones interaction (eq 6) can be estimated from the London equation[63]

$$B_{ij} = \frac{3}{2}\frac{\eta_i\eta_j}{\eta_i + \eta_j}\alpha_i\alpha_j \tag{8}$$

where η and α are the ionization energy and polarizability, respectively, of the subscripted atoms. We follow Tkatchenko[43] in writing the homonuclear B coefficient in the form

$$B_i \propto \eta_i V_i^2 \tag{9}$$

where η_i and V_i are, respectively, the effective ionization energy and volume of atom i. (This assumes that the polarizability is simply proportional to volume;[80] it is worth noting that Gould has argued for an element-specific scaling between these quantities.[78]) As noted above, β^{-1} is related to the atomic radius, so $V_i^2 \propto \beta^{-6}$. For a single atom, the ionization energy is given by $\eta = \frac{\beta^2}{8}$,[61] and we assume the same to be true for the effective ionization energy of an atom in a molecule. This is distinct from the Tkatchenko method, in that we allow each atom to have a unique effective ionization energy, which contributes to the determination of the LJ parameters. Hence, inserting the element-specific fitting parameter C_{B,ϵ_i}, we have

$$B_i = C_{B,\epsilon_i}\eta_i\beta_i^{-6} = \frac{C_{B,\epsilon_i}}{8}\beta_i^{-4} \tag{10}$$

Intuitively, the lower the value of β_i, and hence the more diffuse the electron density, the greater the dispersion coefficient. The element-specific fitting parameters C_{B,ϵ_i} are adjusted with ForceBalance, as detailed in Section 2.5.[79,80]

From the expressions above, we can also derive that

$$\varepsilon_i = \frac{C_{B,\epsilon_i}\eta_i}{4C_{\sigma,\epsilon_i}^{6}} = \frac{C_{B,\epsilon}\beta_i^2}{32C_{\sigma,\epsilon_i}^{6}} \tag{11}$$

Since ε corresponds to the depth of the LJ energy well, this says that a more diffuse electron density corresponds to a smaller well-depth. This trend reflects the fact that a more diffuse electron density also increases σ and thus causes the energy well to be at a greater distance where dispersion forces will be weaker. Note that some prior methods of deriving the dispersion term by AIM methods have neglected variations in the ionization energy[43−45] among atoms. From the present expression for ε, it is apparent that this neglect causes all atoms of a given element to have the same value of ε. In the present approach, different atoms of a given element can have different values of both σ and ε. These issues are further considered in the Discussion section.

2.5. Optimizing the Elemental Mapping Parameters Using ForceBalance. We used ForceBalance,[34] a software package that automatically adjusts fitting parameters using parametric gradients of simulated properties, to optimize the mapping parameters $C_{\sigma,\epsilon}$ and $C_{B,\epsilon}$ for the elements carbon, nitrogen, oxygen, and separately for polar and nonpolar hydrogen. These mapping parameters were adjusted so that simulations of seven pure organic liquids (Section 2.6) with the resulting LJ parameters would yield properties close to

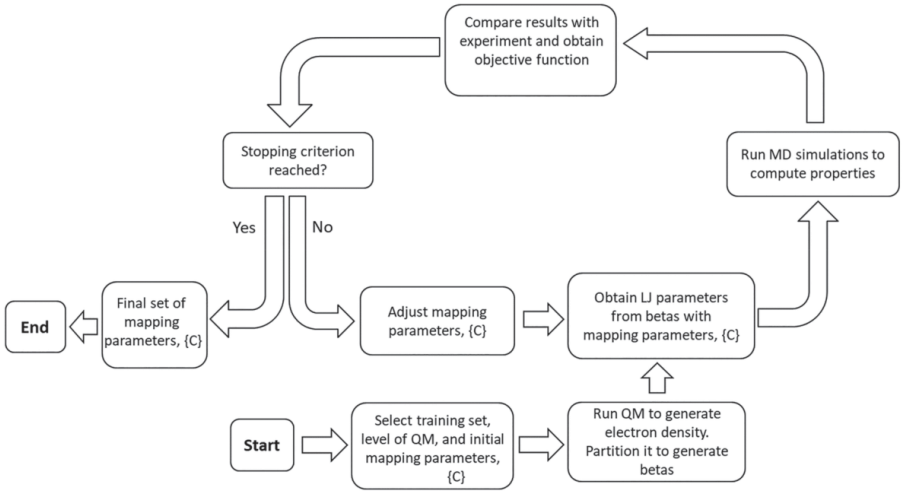


Figure 3. Optimization of the ten QM-to-FF mapping parameters. Quantum mechanical calculations are carried out on training set molecules to compute beta for each atom. These values are mapped to LJ parameters using the mapping parameters. Simulations are carried out to compare simulation results to experiment, and the parameters are iteratively updated based on the gradient of the objective function in parameter space. Once the mapping parameters are optimized, they are saved and used to generate LJ parameters for other molecules.

experiment. The adjusted mapping parameters were then tested by using simulations to compute the properties of 24 pure organic liquids outside the training set and comparing these with experiment. The test- and training-set molecules were chosen to be small and simple, so that simulations could be rapidly converged, while still representing significant chemical diversity. In addition, to test transferability, we included test-set compounds with functional groups absent from the training set. For this initial study, atom-centered partial charges were assigned with the AM1/BCC method,[64] as noted above, and bonded parameters were drawn from GAFF,[65] using the program Antechamber.[66] The iterative ForceBalance process was initiated from a set of mapping parameters that minimize the sum of squared deviations between the mapped LJ parameters and GAFF 1.7 LJ parameters for the training set compounds. The stopping criterion for ForceBalance was essentially chosen manually, as the history of the objective function was evaluated to see if any meaningful improvements were being made. When the objective function had fallen significantly from its starting value and appeared to plateau, the ForceBalance program was halted, and the parameters were evaluated.

For each training set compound, liquid phase simulations were performed with the AMBER molecular dynamics suite[67] to compute the heat of vaporization and density in the NPT ensemble at 298 K and 1 atm. The Berendsen barostat and Langevin thermostat were used for all production simulations, and SHAKE was used to constrain all R−H bond lengths. For each calculation, 1000 molecules were used in the simulation box, and production simulations of 12 ns at 2 fs time steps were run. A cutoff of 8 Å was used for both the Particle Mesh Ewald and Lennard-Jones calculations. The mapping parameters were optimized over multiple ForceBalance iterations so that they produced LJ parameters that minimize a regularized, weighted least-squares objective function computed from the squared deviations of the calculated observables and experimental reference data.[34]

The ForceBalance objective function was described in previous work[34] and is briefly summarized here. It has a hierarchical structure with the top level given by the formula

$$L_{\mathrm{tot}}(\mathbf{k}) = \sum_{T \in \mathrm{targets}} w_T L_T(\mathbf{k}) + w_{\mathrm{reg}} |\mathbf{k}|^2$$

where the total objective function L_{tot} depends on the optimization variables $\mathbf{k}$ and is equal to the sum of contributions from the parametrization targets L_T weighted by w_T, plus a Tikhonov regularization term weighted by w_{reg}. Each parametrization target is a weighted sum of contributions for one or more properties:

$$L_T(\mathbf{k}) = \sum_{j \in \mathrm{properties}} w_j^{(T)} L_j^{(T)}(\mathbf{k})$$

In this study, the target weights w_T and the property weights $w_j^{(T)}$ were set to unity for both properties used (density and heat of vaporization), allowing each to contribute equally. The term for each property $L_j^{(T)}(\mathbf{k})$ is given by a weighted and normalized sum over individual data points

$$L_j^{(T)}(\mathbf{k}) = \frac{1}{(d_j^{(T)})^2} \frac{\sum_{p \in \mathrm{points}} w_{jp}^{(T)} \left| y_{jp}^{(T)}(\mathbf{k}) - y_{jp,\mathrm{ref}}^{(T)} \right|^2}{\sum_{p \in \mathrm{points}} w_{jp}^{(T)}}$$

where $y_{jp}^{(T)}$ and $y_{jp,\mathrm{ref}}^{(T)}$ are, respectively, the simulated and reference data point for property j and point p within target T. The quantity $d_j^{(T)}$ is a scaling factor used to normalize and remove physical units for property j and has the same effect as an inverse square weight; we used values of 30 mg/ml for density and 0.3 kJ/mol for heat of vaporization.

The optimization variables $\mathbf{k}$ are mapped to a set of physical parameters $\mathbf{K}$ by an exponential mapping as $K_\lambda = K_\lambda^{(0)} \exp[k_\lambda]$, where $K_\lambda^{(0)}$ represents the original parameter values. Under this

exponential mapping, the physical parameters do not change sign from their original values. The regularization term may be expressed in terms of mathematical parameters as

$$w_{\mathrm{reg}} |\mathbf{k}|^2 = w_{\mathrm{reg}} \sum_{\lambda \in \mathrm{params}} k_\lambda^2 = w_{\mathrm{reg}} \sum_{\lambda \in \mathrm{params}} \left(\ln \frac{K_\lambda}{K_\lambda^{(0)}} \right)^2$$

During the optimization, FB computes the gradients of simulated properties with respect to force field parameters, i.e., $\nabla_{\mathbf{k}} y_{jp}^{(T)}(\mathbf{k})$, and uses them to construct the gradient and approximate Hessian of L_{tot} in the parameter space. In this work, $y_{jp}^{(T)}(\mathbf{k})$ represents ensemble properties obtained from thermodynamic sampling, and $\nabla_{\mathbf{k}} y_{jp}^{(T)}(\mathbf{k})$ is computed using thermodynamic fluctuation formulas as described in previous work.[68]

The overall computational workflow is diagrammed in Figure 3. Once the mapping parameters were optimized, they were used to generate LJ parameters for a larger test set consisting of 24 molecules, for which densities, heats of vaporization, and heat capacities were calculated and compared with experimental numbers. The uncertainties reported for the calculated properties were obtained by a previously described blocking method.[69,70]

2.6. Training and Test Data. When optimizing and testing force field parameters against experimental observables, it is essential to use reliable data. Here, we obtained data from the ThermoML[71] archive provided by NIST and used a separate compilation of liquid properties of organic molecules[72] as a cross-check to ensure accurate numbers for both the training and test sets. For some values from the ThermoML archive, the values were taken as averages over multiple experimental sources. While experimental uncertainties are not typically provided for these data, the few compounds that do have uncertainties typically show them to be of the order of less than 1% standard deviation (for all properties examined), even when measurements that are nearly a century old[73] are included when calculating average and standard deviations. The experimental uncertainties for these compounds are reported as under ∼5% for both heats of vaporization and densities. The training set comprised the following seven pure liquids: methanol, ethanol, aminoethanol, acetaldehyde, ethylamine, benzene, and acetonitrile (see Figure 4). The test set comprises the 24 additional pure liquids listed

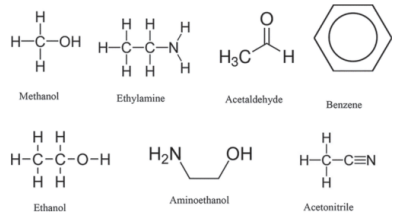


Figure 4. Molecules used to train the elemental mapping parameters.

in Table 3. In order to rigorously assess transferability of the fitted parameters, we chose test-set compounds with functional groups not in the training set. Radial distribution functions were digitized from figures in various sources[74−76] to allow comparison of radial distribution functions calculated from simulations of both SDLJ and GAFF parameters. Radial distribution functions were computed for the present simulations with the cpptraj program.[77]

As the observation may be relevant for other studies that rely on pure liquid properties, it is perhaps worth also reporting that a few compounds initially included in the training or test sets were found to undergo very slow conformational interconversions, no matter what starting conformer was used, even in simulations as long as 10 ns. In particular, we observed few or no syn-anti conversions of the proton in the carboxylic acids formic and acetic acid, either in gas or liquid phases. Both QM calculations and standard force fields point to a barrier between the two states of at least 6 kcal/mol,[78] so we do not think this problem results from the particular parameters used here. Simulations of these molecules led to significant convergence problems in the thermodynamic properties, so these compounds were removed and are not present in this study. Similarly, short esters were removed for the same reason.

3. RESULTS

This section first reports on the optimization of the ten elemental mapping parameters using ForceBalance and a small training set of molecules. Then the transferability of the resulting parameters is tested with a larger, nonoverlapping set of 24 test molecules. The results are compared with experiment and with corresponding simulations using GAFF LJ parameters.

3.1. Optimization of Elemental Mapping Parameters Using ForceBalance. As detailed in Methods, an electronic structure calculation was run for each compound in the training set, and the MBIS method was used to compute β_i for each atom i in each compound. (Final β values for all atoms in each molecule in the training set are provided in Supplementary Table 4.) These quantities were then used with the expressions in Section 2.4, and after about 40 iterations of optimization of the mapping parameters using ForceBalance (Section 2.5), a large improvement in the ForceBalance objective function[34] was observed (Figure 5). The procedure

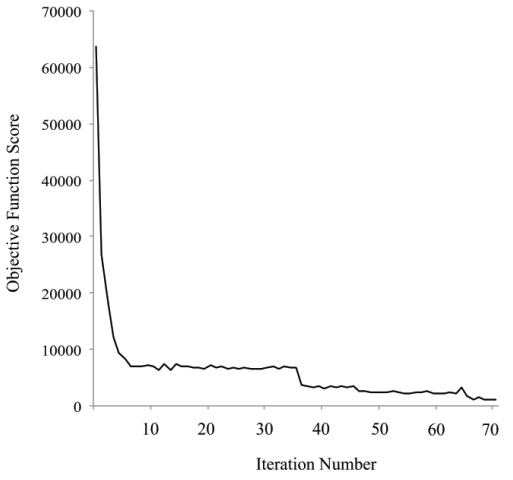


Figure 5. History of objective function of training set over the entirety of the optimization run using ForceBalance.

Table 1. Final Results for Training Set Molecules[a]

	density (mg/ml)			ΔH_{Vap} (kJ/mol)		
compound	experimental	computed	±	experimental	computed	±
methanol	784	762.5	0.50	31.3	36.1	0.28
ethanol	789	784.0	0.55	42.3	42.3	0.68
aminoethanol	1011	1024.8	0.34	58.0	60.2	0.71
ethylamine	688	709.1	0.21	29.0	29.3	0.51
acetaldehyde	784	813.0	0.41	26.1	28.3	0.50
benzene	876	880.9	0.53	33.9	32.6	0.52
acetonitrile	786	747.3	0.40	33.4	31.8	0.52
MUE		**2.4%**			**5.0%**	

[a]Percent mean unsigned errors (MUE) are computed as $100 \sum_i^{N_{data}} \frac{|C_i - E_i|}{C_i}$ where N_{data} is the number of data (compounds), and C_i and E_i are the computed and experimental quantities, respectively, for compound i. Uncertainties were obtained by the blocking method (see Methods).

led to modest additional improvement as it ran out to about 70 cycles. The values of the final Lennard-Jones parameters are provided in Table 2 (in the standard AMBER form $R_{min}/2$ and ε), and the densities and heats of vaporization of the training set molecules computed with the optimized parameters are compared with experiment in Table 1. Final values of the mapping parameters are provided in Supplementary Table 1. Additionally, select mol2 and frcmod files are provided in the GitHub repository for a few test-set molecules.

It is expected that the value of β for each atom in a molecule, and hence the values of σ and ε assigned by this method, will depend to some degree on the conformation used for the QM calculation. To assess the sensitivity to conformation, we took several snapshots of butanol from a 300 K gas-phase simulation and used the trained parameters to compute LJ parameters for all atoms in each conformation. As detailed in Supplementary Tables 2 and 3, the variations across conformations are small, with standard deviations in σ of at most 0.02 Å, and standard deviations in ε of at most 0.001 kcal/mol.

3.2. Test-Set Validation of Optimized Mapping Parameters. The trained SDLJ method yields densities for the test set that agree with experiment about as well as those in the training set and heats of vaporization with about double the relative mean unsigned error of the training set (Figure 6, Table 1, and Table 3). Importantly, the SDLJ parameters provide good agreement with experiment even for compounds with functional groups distinctly different from those in the training set. For example, SDLJ reproduces the properties of dioxolane and furan reasonably well and indeed more closely than done by GAFF, although the training set includes no substituted phenyls or furans. Taken together, these observations suggest that the parameters were not overfitted. Overall, the new method yields accuracy on the test set similar to that obtained with GAFF LJ parameters (Table 3 and Figure 6). This is despite the fact that the SDLJ method has only ten fitting parameters and was trained on only 14 observables. In contrast, the test-set compounds span 16 GAFF atom types and thus include 32 LJ parameters that are, at least in principle, independently adjustable. However, it should also be noted that GAFF was not parametrized against the present training set of liquid properties.

At the same time, it is worth noting that some compounds show undesirably large errors when modeled with either SDLJ or GAFF. Examples include the density of formamide for SDLJ and especially GAFF, the heat of vaporization of propionitrile for SDLJ, and the heat of vaporization of *o*-xylene and butanol for GAFF. Further work is needed to assess whether such

Table 2. Values of $R_{min}/2$ and ε for Training Set Molecules Generated from the Final Optimized Values of the Elemental Mapping Parameters

molecule	atom	$R_{min}/2$ (Å)	ε (kcal/mol)
methanol			
	c	1.936	0.079
	h-c	1.448	0.025
	o	1.486	0.149
	h-o	0.776	0.028
ethanol			
	c1	1.916	0.081
	c2	2.012	0.074
	h-c1	1.439	0.025
	h-c2	1.423	0.025
	o	1.488	0.148
	h-o	0.779	0.028
aminoethanol			
	c-n	1.946	0.079
	c-o	1.939	0.079
	o	1.497	0.146
	n	1.958	0.304
	h-n	0.782	0.027
	h-o	0.755	0.029
	h-c-n	1.412	0.026
	h-c-o	1.447	0.025
ethylamine			
	c1	2.010	0.074
	c2	1.928	0.080
	h-c1	1.426	0.025
	h-c2	1.435	0.025
	n	1.957	0.304
	h-n	0.802	0.026
acetaldehyde			
	c1	1.999	0.075
	c2	1.843	0.088
	h-c1	1.419	0.026
	h-c2	1.423	0.025
	o	1.478	0.150
benzene			
	c	1.972	0.076
	h-c	1.368	0.028
acetonitrile			
	c-c	1.986	0.075
	c-n	1.845	0.088
	h-c-n	1.374	0.027
	n	1.891	0.326

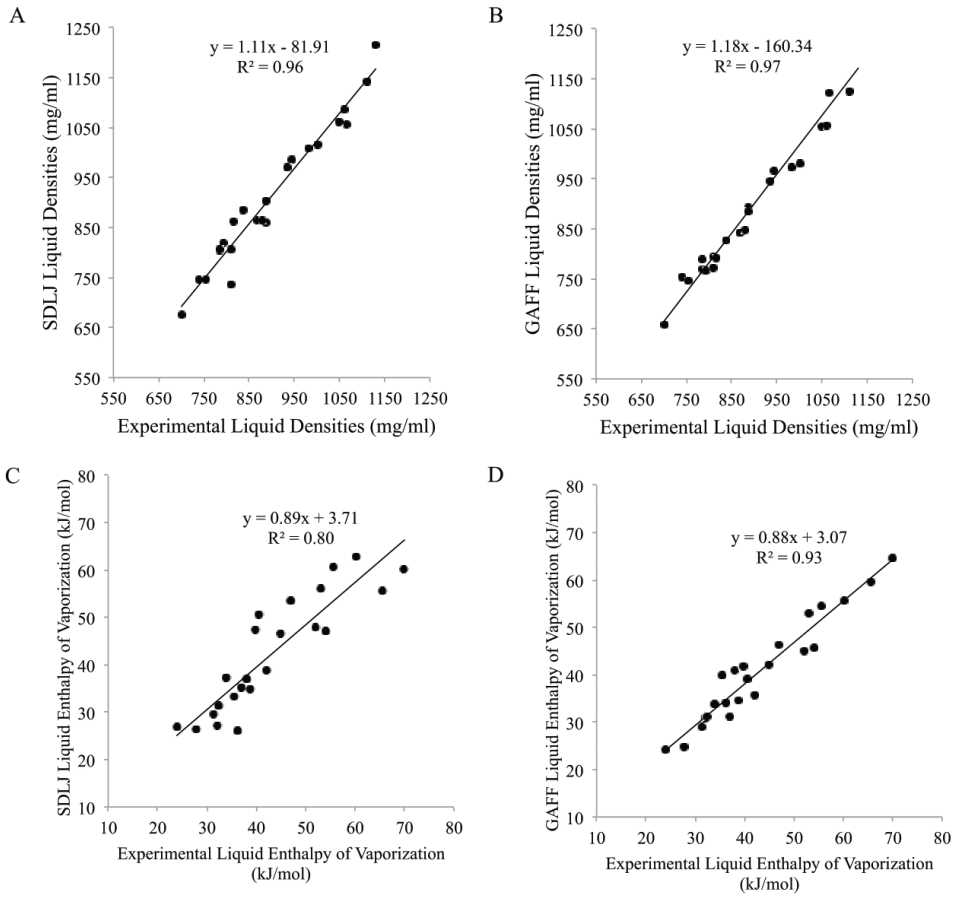


Figure 6. Scatter plots of experimental values of density and heat of vaporization versus SDLJ or GAFF simulated results. (A) Experimental and SDLJ densities. (B) Experimental and GAFF densities. (C) Experimental and SDLJ heats of vaporization. (D) Experimental and GAFF heats of vaporization. Linear regression results are provided on each panel.

errors should be attributed to problems with the LJ parameters; problems with other parameters, such as partial charges; from limitation in the training set; or, perhaps, from limitations of the common functional form itself.

We further probed the reliability of the SDLJ-based FF by computing radial distribution functions (RDFs) and comparing them with available experimental[72−74] data. As shown in Figure 7, SDLJ tends to overestimate the sharpness of the first shell hydration peaks for H-bonding atoms in methanol and ethanol, while GAFF tends to underestimate these first-shell peaks. However, SDLJ does a better job of capturing longer ranged structure in these liquids, such as the subtle valley in the tail of the ethanol O−O interaction. It is perhaps worth noting here that, unlike GAFF, SDLJ assigns polar hydrogens a nonzero radius. For methylamine and benzene (Figure 8), SDLJ does a somewhat better job than GAFF of capturing the overall shape and details of these less peaked RDFs. Overall, SDLJ does a reasonable job of capturing the fine structure of these liquids, even though the method was not trained on these data.

4. DISCUSSION

This study has demonstrated the feasibility of constructing a physics-based, QM-to-FF mapping, which generates LJ parameters that yield pure liquid properties whose accuracy is similar to that of the well-accepted GAFF force field, despite having many fewer adjustable parameters. The transferability of the mapping is supported by the fact that good results were obtained for test-set compounds having functional groups not represented in the training set. This work is founded on important prior advances in AIM analysis,[57] dispersion interactions,[43] and automated parameter optimization.[34]

A key feature of our approach is the abandonment of atom-typing in the assignment of LJ parameters. Instead, under the present schema, each atom of a molecule is assigned unique LJ parameters based on the QM calculation. This is advantageous, as it largely sidesteps the challenge of categorizing atoms according to their chemical environment. Indeed, although the atom type categorizations used in today's FFs are useful and are well-motivated by chemical logic, it is not clear that they represent an optimal balance between parsimony and accuracy,

Table 3. Test-Set Results for the SDLJ Method and GAFF[a]

	density (mg/ml)					ΔH_{VAP} (kJ/mol)				
compound	exp.	SDLJ	±	GAFF	±	exp.	SDLJ	±	GAFF	±
propyl acetate	888	903.8	0.29	892.3	0.44	38	37.1	0.59	41.1	0.62
acetone	784	805.5	0.34	769.0	0.28	31.3	29.5	0.44	29.0	0.45
THF	889	859.9	0.42	884.5	0.32	32.2	27.1	0.51	30.9	0.56
DMF	944	987.5	0.27	965.7	0.54	46.8	53.6	0.51	46.4	0.57
propionitrile	792	818.6	0.48	767.4	0.29	36.3	26.1	0.41	34.0	0.42
methylamine	700	677.3	0.29	658.9	0.51	23.8	26.8	0.38	24.2	0.31
butylamine	740	745.6	0.24	753.6	0.45	34	37.3	0.62	33.9	0.59
butanol	810	735.7	0.49	794.8	0.29	52	48.0	0.71	45.0	0.68
isopropyl alcohol	786	808.2	0.29	790.0	0.35	45	46.5	0.43	42.2	0.59
glycol	1110	1143.8	0.47	1125.2	0.39	65.6	55.7	0.55	59.7	0.43
phenyl-2-propanone	1001	1015.6	0.41	981.7	0.25	55.5	60.6	0.68	54.6	0.73
furan	936	971.0	0.42	946.4	0.41	27.7	26.4	0.46	24.9	0.36
formamide	1130	1214.5	0.39	1260.9	0.96	60.2	62.9	0.37	55.7	0.37
propenal	839	884.9	0.43	828.6	0.34	32.3	31.4	0.34	31.1	0.41
dioxolane	1065	1057.3	0.40	1123.6	0.41	35.5	33.3	0.50	40.0	0.44
propenoic acid	1050	1063.0	0.44	1055.5	0.32	53.1	56.2	0.59	52.9	0.55
toluene	867	865.5	0.32	843.3	0.34	37	35.1	0.52	31.1	0.54
1,3-propanediol	1060	1086.3	1.72	1056.4	0.54	70	60.1	0.59	64.8	1.09
3-pentanone	809	807.6	0.38	772.7	0.31	38.7	34.9	0.66	34.5	0.54
o-xylene	880	864.8	0.28	848.1	0.61	42	38.9	0.64	35.8	2.85
pyridine	982	1009.8	0.49	974.8	0.40	40.5	50.5	0.51	39.2	0.42
pentylamine	755	746	0.39	745.9	0.36	39.7	47.5	0.62	41.8	0.56
3-pentanol	815	863.2	0.38	793.0	0.29	54	47.0	0.70	45.9	0.63
%MUE		**3.0%**		**2.5%**			**11%**		**7.8%**	
MSE		**13.8**		**−0.1**			**−0.8**		**−2.2**	

[a]See Table 1 for definitions. Uncertainties were obtained by the blocking method (see Methods).

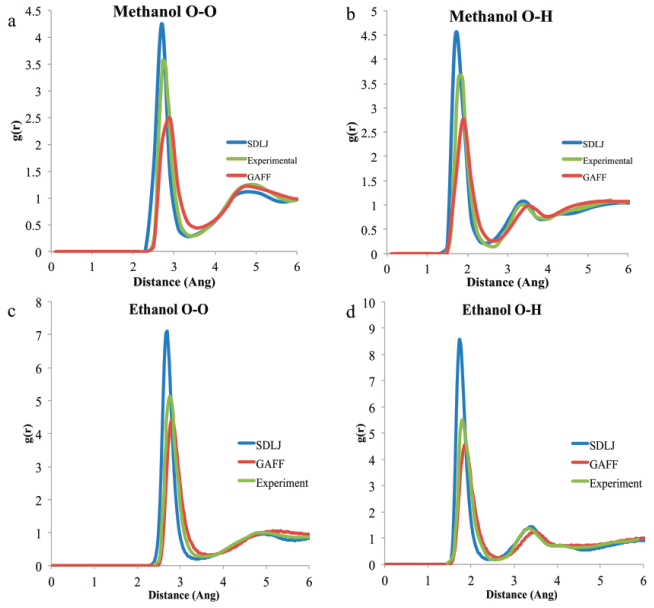


Figure 7. Radial distribution functions of various pair interactions in neat methanol and ethanol.

and the requirement for atom-typing has been cited as a problematic aspect of FF parametrization.[42,76] Recently, this problem has been addressed with a demonstration that the typing itself, rather than just the parameters associated with a fixed set of types, can be sampled effectively.[36] Here, we have considered a second approach, one which does away entirely

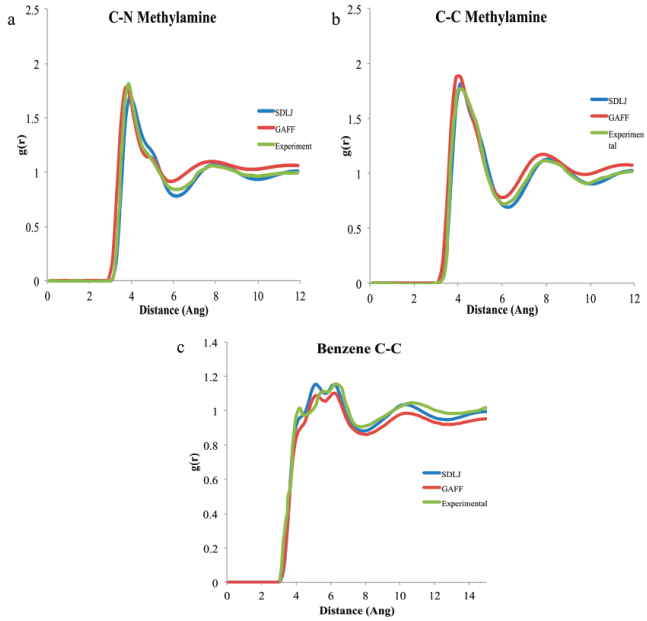


Figure 8. Radial distribution functions of various pair interactions in neat methylamine and benzene.

with LJ types. Further work is needed to ascertain which, if either, of these broad approaches will be most effective. An advantage of atom-typing is that it allows *ad hoc* adjustments that may at times be helpful. On the other hand, the present approach is advantageous in that it can facilitate comprehensive parameter optimization or Bayesian sampling by reducing the number of adjustable parameters and removes the need to sample over the typing itself. It is perhaps encouraging that one of the leading methods of assigning atomic partial charges, namely RESP, similarly eschews typing and instead uses a purely physics-based QM-to-FF mapping.

The adjustment of the mapping parameters presumably allows them to capture or compensate for several issues in the primary QM calculation and the physical model used for the mapping. First, although the AIM concept is intuitively pleasing, it is at best a physical approximation, so it is probably inevitable that some adjustment is needed. Second, even if an AIM analysis could provide flawless LJ parameters, adjustment would be needed to compensate for deficiencies in other FF terms, such as charge−charge interactions, for complexities that arise on going from gas to condensed phase, such as many-body effects, and for the neglect of nuclear quantum mechanics in typical classical simulations. By the same token, although gas-phase QM interaction data may be used to guide the adjustment of FF parameters, these interactions will inevitably change in subtle ways upon going into the condensed phase. We therefore chose to omit any gas-phase QM data in the actual training of the mapping parameters, opting instead to use only condensed phase experimental observables.

Important recent studies have also used a tuned QM-to-FF mapping to assign LJ parameters without atom-typing.[44,45] The present approach is different in two key respects. The first difference is that we have modified only the LJ term and applied the GAFF and AM1/BCC partial charges without change, rather than simultaneously refitting bonded terms and adding off-atom partial charges. This approach makes it possible to isolate the effect of this one change on the accuracy of the FF and also maintains the common functional form and thus compatibility with widely used simulation packages. It is worth noting that all such methods are expected to generate parameters that depend to some degree on the conformation of the molecule used in the QM calculations. In the present case, at least, initial testing shows only an encouragingly small dependence on conformation. It should nonetheless be noted that, if cases are encountered with the dependence on conformation is nontrivial, it should be possible to address these by averaging over thermodynamically accessible conformations and/or using molecular fragmentation approaches so that parameters can be assigned to relatively rigid molecular components.

The second difference is in the QM-to-FF mapping itself. In particular, prior studies have used the Tkatchenko-Scheffler (TS) approach,[43] in which the AIM volume of each atom yields its AIM polarizability,[77] which is used in turn to determine the coefficient of the dispersion interaction (B_i in our notation). Although the London dispersion interaction is determined by not only the polarizability but also the ionization energy,[61] the TS approach assumes, in effect, that any effect of variations in the effective ionization energy across atoms on dispersion interactions is ultimately canceled by other factors. As noted here (Section 2.4.2), this assumption causes all atoms of a given element (e.g., all carbons) to be assigned the identical value of epsilon and thus the same depth of the LJ energy well. Although this approach has led to methods of deriving nonbonded parameters which can provide

good agreement with experiment,[41,42,57] it runs counter to the empirical knowledge of typed force fields which allow both sigma and epsilon to vary for a given element. The present approach thus uses additional physical reasoning to extract not only the atomic volume but also an effective AIM ionization energy for each atom and thus allows for variation in both sigma and epsilon across atoms of a given element.

As discussed above, the great potential benefit of the general approach taken here is the reduction in the dimensionality of the space of LJ parameters that need to be optimized or sampled. This is particularly important when evaluation of the objective function requires running simulations, as this causes each iteration to be computationally costly and thus increases the risk of missing key sectors of LJ parameter space. Thus, we envision applying the SDLJ approach within a broad FF optimization scheme, where the decreased dimensionality will make it easier to carry out a more thorough parameter optimization. We also anticipate using QM-to-FF mappings to generate other parameter types, such as torsional potentials. Additionally, we are working on a version of RESP charges which includes one adjustable parameter that scales the overall polarity of a molecule.[81] Together, these efforts could be combined in a coordinated optimization of an entire force field with only a small set of parameters that need to be optimized against experimental observables. The relatively thorough sampling of parameter space this enables should not only lead to better parameters but also help to assess with greater confidence whether a proposed improvement in the functional form truly enables more accurate simulations than the starting functional form. Thus, the present approach is supportive of a systematic approach to advancing both the parametrization and the form of future force fields.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jctc.9b00713.

Supplementary Table 1, elemental mapping constants; Supplementary Table 2, Lennard-Jones radii for series of conformers of butanol; Supplementary Table 3, Lennard-Jones well depths for a series of conformers of butanol; Supplementary Figure 1, convergence of ForceBalance optimization from different parameter starting points; Supplementary Table 4, beta values for atoms in training set; and Supplementary Figure 2, details of compounds in test set (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Michael K. Gilson − *University of California San Diego, La Jolla, California;* orcid.org/0000-0002-3375-1738; Email: mgilson@ucsd.edu

Other Authors

Sophie M. Kantonen − *University of California San Diego, La Jolla, California;* orcid.org/0000-0002-5001-836X

Hari S. Muddana − *University of California San Diego, La Jolla, California, and OpenEye Scientific Software, Inc., Santa Fe, New Mexico*

Michael Schauperl − *University of California San Diego, La Jolla, California;* orcid.org/0000-0001-5648-8170

Niel M. Henriksen − *University of California San Diego, La Jolla, California, and AtomWise, Inc., San Francisco, California;* orcid.org/0000-0002-7916-0757

Lee-Ping Wang − *University of California Davis, Davis, California;* orcid.org/0000-0003-3072-9946

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jctc.9b00713

Notes

The contents of this paper are solely the responsibility of the authors and do not necessarily represent the official views of the NIH.

The authors declare the following competing financial interest(s): M.K.G. has an equity interest in and is a cofounder and scientific advisor of VeraChem LLC.

■ ACKNOWLEDGMENTS

M.K.G. acknowledges funding from the National Institute of General Medical Sciences (grant number GM061300). L.P.W. acknowledges funding from the National Institute of Allergy and Infectious Diseases (grant number AI130684). M.S. acknowledges support of the Austrian Science Fund (FWF): Erwin Schroedinger fellowship J-4150.

■ REFERENCES

(1) Lindorff-Larsen, K.; et al. Systematic Validation Of Protein Force Fields Against Experimental Data. *PLoS One* **2012**, *7*, e32131.

(2) Durrant, J. D.; Mccammon, J. A. Molecular Dynamics Simulations And Drug Discovery. *BMC Biol.* **2011**, *9*, 71.

(3) Kaminski, G.; Duffy, E. M.; Matsui, T.; Jorgensen, W. L. Free Energies Of Hydration And Pure Liquid Properties Of Hydrocarbons From The OPLS All-Atom Model. *J. Phys. Chem.* **1994**, *98*, 13077−13082.

(4) Lee, T.; et al. GPU-Accelerated Molecular Dynamics And Free Energy Methods In Amber18: Performance Enhancements And New Features. *J. Chem. Inf. Model.* **2018**, *58* (10), 2043−2050.

(5) Salomon-Ferrer, R.; Götz, A. W.; Poole, D.; Le Grand, S.; Walker, R. C. Routine Microsecond Molecular Dynamics Simulations With AMBER On Gpus. 2. Explicit Solvent Particle Mesh Ewald. *J. Chem. Theory Comput.* **2013**, *9*, 3878−3888.

(6) Fine, R.; Dimmler, G.; Levinthal, C. FASTRUN: A Special Purpose, Hardwired Computer For Molecular Simulation. *Proteins: Struct., Funct., Genet.* **1991**, *11*, 242−253.

(7) Ito, T.; et al. A Special-Purpose Computer For Molecular Dynamics: GRAPE-2A. *Proteins: Struct., Funct., Genet.* **1994**, *20*, 139−148.

(8) Shaw, D. E.; et al. Anton 2: Raising The Bar For Performance And Programmability In A Special-Purpose Molecular Dynamics Supercomputer. *SC14: International Conference For High Performance Computing, Networking, Storage And Analysis* **2014**, 41−53.

(9) Shaw, D. E. et al. Millisecond-Scale Molecular Dynamics Simulations On Anton. *SC '09*; 2009; Article No. 39, DOI: 10.1145/1654059.1654099.

(10) Barducci, A.; Bonomi, M.; Parrinello, M. Metadynamics. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2011**, *1*, 826−843.

(11) Shirts, M. R.; Chodera, J. D. Statistically Optimal Analysis Of Samples From Multiple Equilibrium States. *J. Chem. Phys.* **2008**, *129*, 124105.

(12) Sugita, Y.; Okamoto, Y. Replica-Exchange Molecular Dynamics Method For Protein Folding. *Chem. Phys. Lett.* **1999**, *314*, 141−151.

(13) Miao, Y.; Feher, V. A.; Mccammon, J. A. Gaussian Accelerated Molecular Dynamics: Unconstrained Enhanced Sampling And Free

Energy Calculation. *J. Chem. Theory Comput.* **2015**, *11* (8), 3584−3595.
(14) Zheng, L.; Chen, M.; Yang, W. Random Walk In Orthogonal Space To Achieve Efficient Free-Energy Simulation Of Complex Systems. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 20227−20232.
(15) Muddana, H. S.; Fenley, A. T.; Mobley, D. L.; Gilson, M. K. The SAMPL4 Host-Guest Blind Prediction Challenge: An Overview. *J. Comput.-Aided Mol. Des.* **2014**, *28*, 305−317.
(16) Muddana, H. S.; et al. Blind Prediction Of Host-Guest Binding Affinities: A New SAMPL3 Challenge. *J. Comput.-Aided Mol. Des.* **2012**, *26*, 475−487.
(17) Bosisio, S.; Mey, A. S. J. S.; Michel, J. Blinded Predictions Of Host-Guest Standard Free Energies Of Binding In The SAMPL5 Challenge. *J. Comput.-Aided Mol. Des.* **2017**, *31*, 61−70.
(18) Yin, J.; Henriksen, N. M.; Slochower, D. R.; Gilson, M. K. The SAMPL5 Host−Guest Challenge: Computing Binding Free Energies And Enthalpies From Explicit Solvent Simulations By The Attach-Pull-Release (APR) Method. *J. Comput.-Aided Mol. Des.* **2017**, *31*, 133−145.
(19) Geballe, M. T.; et al. The SAMPL2 Blind Prediction Challenge: Introduction And Overview. *J. Comput.-Aided Mol. Des.* **2010**, *24* (4), 259−79.
(20) Skillman, A. G. SAMPL3: Blinded Prediction Of Host−Guest Binding Affinities, Hydration Free Energies, And Trypsin Inhibitors. *J. Comput.-Aided Mol. Des.* **2012**, *26*, 473−474.
(21) Nicholls, A.; et al. Predicting Small-Molecule Solvation Free Energies: An Informal Blind Test For Computational Chemistry. *J. Med. Chem.* **2008**, *51*, 769−779.
(22) Nerenberg, P. S.; Head-Gordon, T. New Developments In Force Fields For Biomolecular Simulations. *Curr. Opin. Struct. Biol.* **2018**, *49*, 129−138.
(23) Leach, A. R. *Molecular Modelling: Principles And Applications*; Prentice Hall: Harlow, England, 2001.
(24) Anisimov, V. M.; et al. Determination Of Electrostatic Parameters For A Polarizable Force Field Based On The Classical Drude Oscillator. *J. Chem. Theory Comput.* **2005**, *1*, 153−168.
(25) Maple, J. R.; et al. A Polarizable Force Field And Continuum Solvation Methodology For Modeling Of Protein−Ligand Interactions. *J. Chem. Theory Comput.* **2005**, *1*, 694−715.
(26) Ponder, J. W.; et al. Current Status Of The AMOEBA Polarizable Force Field. *J. Phys. Chem. B* **2010**, *114*, 2549−2564.
(27) Wang, J.; Cieplak, P.; Luo, R.; Duan, Y. Development Of Polarizable Gaussian Model For Molecular Mechanical Calculations I: Atomic Polarizability Parameterization To Reproduce Ab Initio Anisotropy. *J. Chem. Theory Comput.* **2019**, *15*, 1146−1158.
(28) Van Vleet, M. J.; Misquitta, A. J.; Schmidt, J. R. New Angles On Standard Force Fields: Toward A General Approach For Treating Atomic-Level Anisotropy. *J. Chem. Theory Comput.* **2018**, *14* (2), 739−758.
(29) Freitag, M. A.; Gordon, M. S.; Jensen, J. H.; Stevens, W. J. Evaluation Of Charge Penetration Between Distributed Multipolar Expansions. *J. Chem. Phys.* **2000**, *112*, 7300−7306.
(30) Gootz, T. D.; Subashi, T. A.; Lindner, D. L. Simple Spectrophotometric Assay For Measuring Protein Binding Of Penem Antibiotics To Human Serum. *Antimicrob. Agents Chemother.* **1988**, *32*, 159−163.
(31) Rackers, J. A.; et al. An Optimized Charge Penetration Model For Use With The AMOEBA Force Field. *Phys. Chem. Chem. Phys.* **2017**, *19*, 276−291.
(32) Zanette, C.; et al. Toward Learned Chemical Perception Of Force Field Typing Rules. *J. Chem. Theory Comput.* **2019**, *15* (1), 402−423.
(33) Betz, R. M.; Walker, R. C. Paramfit: Automated Optimization Of Force Field Parameters For Molecular Dynamics Simulations. *J. Comput. Chem.* **2015**, *36*, 79−87.
(34) Wang, L.-P.; Martinez, T. J.; Pande, V. S. Building Force Fields: An Automatic, Systematic, And Reproducible Approach. *J. Phys. Chem. Lett.* **2014**, *5*, 1885−1891.
(35) Sen, G. et al. Comparing Optimization Strategies For Force Field Parameterization. 2018, https://arxiv.org/abs/1812.00326 (accessed Jan 14, 2020).
(36) Brommer, P.; Gähler, F. Potfit: Effective Potentials From Ab-Initio Data. *Modell. Simul. Mater. Sci. Eng.* **2007**, *15*, 295−304.
(37) Hülsmann, M. et al. Optimizing Molecular Models Through Force-Field Parameterization Via The Efficient Combination Of Modular Program Packages. In *Foundations Of Molecular Modeling And Simulation: Select Papers From FOMMS 2015*; Snurr, R. Q., Adjiman, C. S., Kofke, D. A., Eds.; Springer: 2016; pp 53−77, DOI: 10.1007/978-981-10-1128-3_4.
(38) Faller, R.; Schmitz, H.; Biermann, O.; Müller-Plathe, F. Automatic Parameterization Of Force Fields For Liquids By Simplex Optimization. *J. Comput. Chem.* **1999**, *20*, 1009−1017.
(39) Mobley, D. L.; et al. Escaping Atom Types In Force Fields Using Direct Chemical Perception. *J. Chem. Theory Comput.* **2018**, *14*, 6076−6092.
(40) Ge, Y.; Voelz, V. A. Model Selection Using Biceps: A Bayesian Approach For Force Field Validation And Parameterization. *J. Phys. Chem. B* **2018**, *122*, 5610−5622.
(41) Wu, S.; Angelikopoulos, P.; Papadimitriou, C.; Moser, R.; Koumoutsakos, P. A Hierarchical Bayesian Framework For Force Field Selection In Molecular Dynamics Simulations. *Philos. Trans. R. Soc., A* **2016**, *374*, 20150032.
(42) Dutta, R.; Brotzakis, Z. F.; Mira, A. Bayesian Calibration Of Force-Fields From Experimental Data: TIP4P Water. *J. Chem. Phys.* **2018**, *149*, 154110.
(43) Tkatchenko, A.; Scheffler, M. Accurate Molecular Van Der Waals Interactions From Ground-State Electron Density And Free-Atom Reference Data. *Phys. Rev. Lett.* **2009**, *102*, 073005.
(44) Cole, D. J.; Vilseck, J. Z.; Tirado-Rives, J.; Payne, M. C.; Jorgensen, W. L. Biomolecular Force Field Parameterization Via Atoms-In-Molecule Electron Density Partitioning. *J. Chem. Theory Comput.* **2016**, *12*, 2312−2323.
(45) Horton, J. T.; Allen, A. E. A.; Dodda, L. S.; Cole, D. J. Qubekit: Automating The Derivation Of Force Field Parameters From Quantum Mechanics. *J. Chem. Inf. Model.* **2019**, *59*, 1366−1381.
(46) Vandenbrande, S.; Waroquier, M.; Speybroeck, V.; Verstraelen, T. The Monomer Electron Density Force Field (MEDFF): A Physically Inspired Model For Noncovalent Interactions. *J. Chem. Theory Comput.* **2017**, *13*, 161.
(47) Swart, M.; Duijnen, P. T. V. Atomic Radii In Molecules For Use In A Polarizable Force Field. *Int. J. Quantum Chem.* **2011**, *111*, 1763−1772.
(48) Visscher, K. M.; Geerke, D. P. Deriving Force-Field Parameters From First Principles Using A Polarizable And Higher Order Dispersion Model. *J. Chem. Theory Comput.* **2019**, *15*, 1875−1883.
(49) Pérez De La Luz, A.; Aguilar-Pineda, J. A.; Méndez-Bermúdez, J. G.; Alejandre, J. Force Field Parametrization From The Hirshfeld Molecular Electronic Density. *J. Chem. Theory Comput.* **2018**, *14*, 5949−5958.
(50) Mei, Y.; et al. Numerical Study On The Partitioning Of The Molecular Polarizability Into Fluctuating Charge And Induced Atomic Dipole Contributions. *J. Phys. Chem. A* **2015**, *119*, 5865−5882.
(51) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development And Testing Of A General Amber Force Field. *J. Comput. Chem.* **2004**, *25*, 1157−1174.
(52) Hanwell, M. D.; et al. Avogadro: An Advanced Semantic Chemical Editor, Visualization, And Analysis Platform. *J. Cheminf.* **2012**, *4*, 17.
(53) Frisch, M. et al. *Gaussian 09*, Revision B01; Gaussian, Inc.: Wallingford, CT, 2009.
(54) Kummel, H. A Biography Of The Coupled Cluster Method *Recent Profess In Many Body Theories. Int. J. Mod. Phys. B* **2003**, *17*, 5311−5325.
(55) Lebedev, V. I. Values Of The Nodes And Weights Of Ninth To Seventeenth Order Gauss-Markov Quadrature Formulae Invariant Under The Octahedron Group With Inversion. *USSR Comput. Math. Math. Phys.* **1975**, *15*, 44−51.

(56) Manz, T. A.; Limasa, N. G. Introducing DDEC6 Atomic Population Analysis: Part 1. Charge Partitioning Theory And Methodology. *RSC Adv.* **2016**, *6*, 47771−47801.
(57) Heidar-Zadeh, F.; Ayers, P. W. How Pervasive Is The Hirshfeld Partitioning? *J. Chem. Phys.* **2015**, *142*, 044107.
(58) Verstraelen, T.; et al. Hirshfeld-E Partitioning: AIM Charges With An Improved Trade-Off Between Robustness And Accurate Electrostatics. *J. Chem. Theory Comput.* **2013**, *9* (5), 2221−2225.
(59) Verstraelen, T.; et al. Minimal Basis Iterative Stockholder: Atoms In Molecules For Force-Field Development. *J. Chem. Theory Comput.* **2016**, *12*, 3894−3912.
(60) Hoffmann-Ostenhof, M.; Hoffmann-Ostenhof, T. 'Schrodinger Inequalities" And Asymptotic Behavior Of The Electron Density Of Atoms And Molecules. *Phys. Rev. A: At., Mol., Opt. Phys.* **1977**, *16*, 1782−1785.
(61) Van Vleet, M. J.; Misquitta, A. J.; Stone, A. J.; Schmidt, J. R. Beyond Born-Mayer: Improved Models For Short-Range Repulsion In Ab Initio Force Fields. *J. Chem. Theory Comput.* **2016**, *12*, 3851−3870.
(62) Lorentz, H. A. Ueber Die Anwendung Des Satzes Vom Virial In Der Kinetischen Theorie Der Gase - Lorentz. *Annalen Der Physik* **1881**, *248*, 127−136.
(63) Eisenschitz, R.; London, F. Über Das Verhältnis Der Van Der Waalsschen Kräfte Zu Den Homöopolaren Bindungskräften. *Eur. Phys. J. A* **1930**, *60*, 491−527.
(64) Jakalian, A.; Jack, D.; Bayly, C. Fast, Efficient Generation Of High Quality Atomic Charges. AM1 - BCC Model: II. Parameterization And Validation. *J. Comput. Chem.* **2002**, *23*, 1623−41.
(65) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development And Testing Of A General Amber Force Field. *J. Comput. Chem.* **2004**, *25*, 1157−1174.
(66) Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A. Automatic Atom Type And Bond Type Perception In Molecular Mechanical Calculations. *J. Mol. Graphics Modell.* **2006**, *25*, 247−260.
(67) Salomon-Ferrer, R.; Case, D. A.; Walker, R. C. An Overview Of The Amber Biomolecular Simulation Package. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2013**, *3*, 198−210.
(68) Wang, L.-P.; et al. Systematic Improvement Of A Classical Molecular Model Of Water. *J. Phys. Chem. B* **2013**, *117*, 9956−9972.
(69) Flyvbjerg, H.; Petersen, H. G. Error Estimates On Averages Of Correlated Data. *J. Chem. Phys.* **1989**, *91*, 461−466.
(70) Henriksen, N. M.; Fenley, A. T.; Gilson, M. K. Computational Calorimetry: High-Precision Calculation Of Host−Guest Binding Thermodynamics. *J. Chem. Theory Comput.* **2015**, *11*, 4377−4394.
(71) Frenkel, M.; et al. Thermomlan XML-Based Approach For Storage And Exchange Of Experimental And Critically Evaluated Thermophysical And Thermochemical Property Data. 1. Experimental Data. *J. Chem. Eng. Data* **2003**, *48*, 2−13.
(72) *Enthalpies Of Vaporization Of Organic Compounds: A Critical Review And Data Compilation*; Blackwell Scientific: 1985.
(73) Mathews, J. H. The Accurate Measurement Of Heats Of Vaporization Of Liquids. *J. Am. Chem. Soc.* **1926**, *48* (3), 562−576.
(74) Figueroa-Gerstenmaier, S.; Giudice, S.; Cavallo, L.; Milano, G. A Molecular Model For H2 Interactions In Aliphatic And Aromatic Hydrocarbons. *Phys. Chem. Chem. Phys.* **2009**, *11*, 3935−42.
(75) Kosztolányi, T.; Bakó, I.; Palinkas, G. Hydrogen Bonding In Liquid Methanol, Methylamine, And Methanethiol Studied By Molecular-Dynamics Simulations. *J. Chem. Phys.* **2003**, *118*, 4546−4555.
(76) Saiz, L.; Padró, J. A.; Guàrdia, E. Structure And Dynamics Of Liquid Ethanol. *J. Phys. Chem. B* **1997**, *101*, 78−86.
(77) Roe, D. R.; Cheatham, T. E. *J. Chem. Theory Comput.* **2013**, *9*, 3084−3095.
(78) Lim, V. T.; et al. Assessing The Conformational Equilibrium Of Carboxylic Acid Via Quantum Mechanical And Molecular Dynamics Studies On Acetic Acid | Journal Of Chemical Information And Modeling. *J. Chem. Inf. Model.* **2019**, *59* (5), 1957−1964.
(79) Grimme, S. A General Quantum Mechanically Derived Force Field (QMDFF) For Molecules And Condensed Phase Simulations. *J. Chem. Theory Comput.* **2014**, *10*, 4497−4514.
(80) Brinck, T.; Murray, J. S.; Politzer, P. Polarizability And Volume. *J. Chem. Phys.* **1993**, *98*, 4305−4306.
(81) Schauperl, M. et al. Force Field Partial Charges With Restrained Electrostatic Potential 2 (RESP2). 2019, DOI: 10.26434/Chemrxiv.10072799.V1.

Chapter 2, in full, is a reprint of the material as it appears in “Data-Driven Mapping of Gas-Phase Quantum Calculations to General Force Field Lennard-Jones Parameters”, SM Kantonen**,** H Muddana, M Schauperl, NM Henriksen, LP Wang, MK Gilson. Journal of Chemical Theory and Computation (2019). The dissertation author was primary investigator and author of this publication.

3 SINGLE SITE FUNCTIONALIZATION OF BETA-CYCLODEXTRIN

3.1 HOST-GUEST CHEMISTRY

Host-guest chemistry describes the interaction between two or more molecules to form multimeric complexes[1]. These complexes are held together by non-bonded interactions, ranging from ionic to hydrophobic, and can be incredibly strong. For example, the dissociation constants of certain adamantyl compounds and the host molecule cucurbit[8]uril are in the picomolar range, rivaling even some of the highest-affinity protein-ligand complexes[2–4]. This capability, coupled with the smaller size of host-guest complexes, makes them ideal candidates for simulations to test binding calculation accuracy.

There are a number of soluble and synthesizable host molecules, which are potential candidates for using as models for binding. Popular choices have included the large cavitand molecule octa-acid[5], the cucurbituril family[2–4], and the cyclodextrin family[6,7]. Each of these molecules contain a large hydrophobic cavity that allows smaller molecules to fit into. Some clever host designs even incorporate decorations alongside the top to restrict the size of the binding partners[8,9], or modulate solubility[5]. In terms of binding proclivity, the cyclodextrin family is known to bind a very wide variety of molecules[6], and the chemical structure provides some biologically relevant functional groups. Additionally, cyclodextrins are known to be quite soluble, owing to the number of exposed hydroxyls, and native cyclodextrins are readily available and inexpensive starting materials . For these reasons, we chose to focus on cyclodextrins as hosts to explore in terms of modifying.

3.2 β-CYCLODEXTRIN

As mentioned above, cyclodextrin is often times a popular choice due to its solubility and proclivity as a binder. There are three main varieties of cyclodextrin, all

made up of varying numbers of glucose monomers, ranging from six to eight, and termed alpha, beta, and gamma cyclodextrin respectively. The number of glucose monomers present in the cyclodextrin influence the size of the cavity, and thus binding.

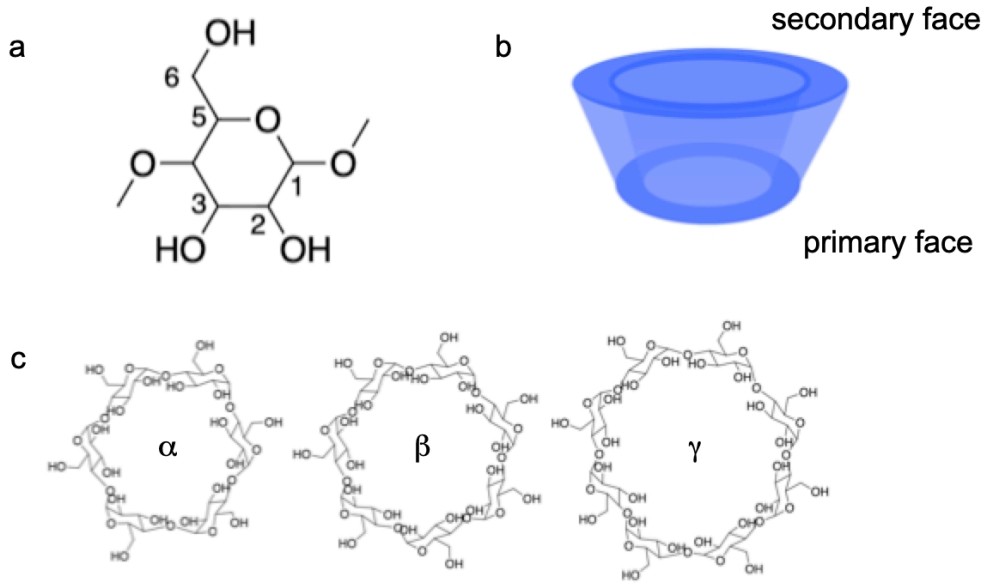


Figure 3.1: Illustrations of various parts of cyclodextrin. (a) the glucose monomer component of cyclodextrin (b) a 3d representation of the general shape of cyclodextrins (c) specific representations of each of the isomers of cyclodextrin.

For this work, we chose to focus on β-CD, which contains seven glucose subunits. This provides a cavity large enough for an adamantyl compound[10], as well as a number of drug molecules[11], to sit nicely into the hydrophobic cavity. Indeed, we performed isothermal titration calorimetry experiments on two cationic adamantyl compounds, amantadine and rimantadine, and saw reasonable affinities for both of the guest molecules (Figure 3.2).

We also reasoned that, in future work, adamantyl compounds could be similarly modified, such that the hydrophobic core would remain constant (and would thus always sit in the cavity), but the top portion could be modified so that it could interact with a modification on cyclodextrin.

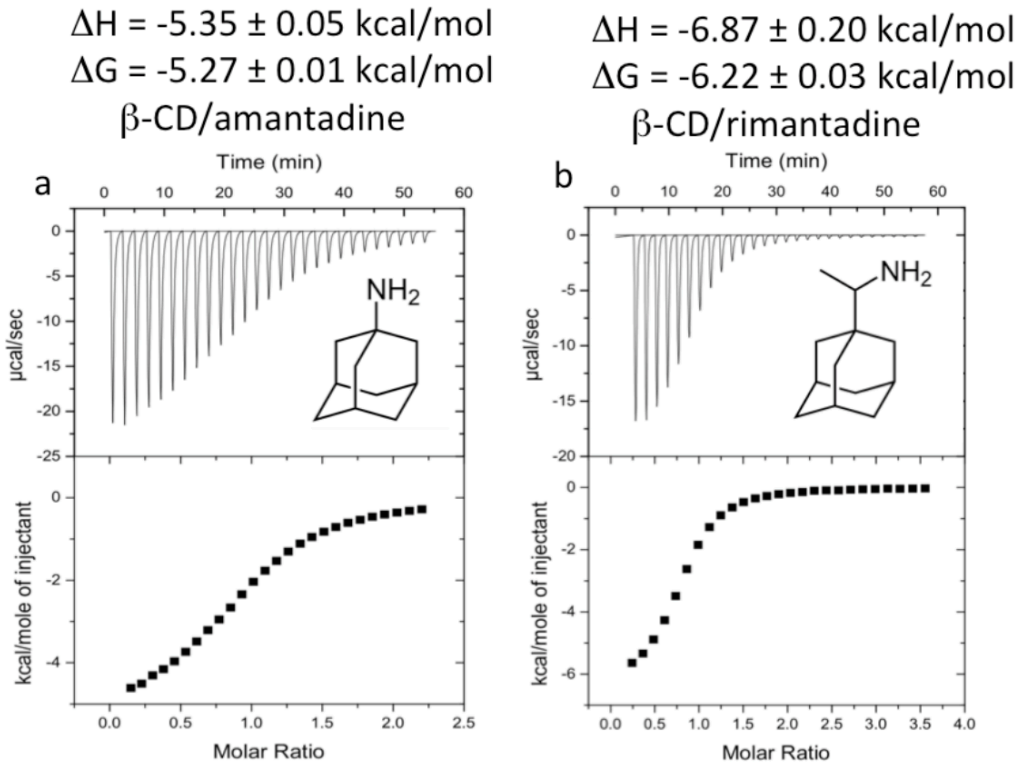


Figure 3.2: Binding of amantadine and rimantadine to beta-cyclodextrin. (a) binding of amantadine to beta-cyclodextrin (b) binding of rimantadine to beta-cyclodextrin.

3.3 SINGLE SITE MODIFICATION OF β-CYCLODEXTRIN

Once we selected β-cyclodextrin to use as a host molecule, we sought a method that would allow us to modify it in such a way that was controllable and gave a pure result. Any racemic mixture or random modification makes comparison *in silico* problematic. Additionally, we wanted to start by modifying the secondary face (see Figure 3.1), as our initial guests, amantadine and rimantadine, bind with their functional

groups sticking outside of the secondary face. Thus, the two most obvious choices to attempt to modify were the 2 and 3 position hydroxyls.

One plausible route is an ether synthesis; this would essentially involve reacting an alkyl halide with a deprotonated secondary face hydroxyl. However, preserving selectivity and containing the degree of functionalization would be difficult to control. We therefore turned to a report by Casas-Solvas[12] showing that a single site modification at the 2-hydroxyl was obtainable by reacting b-cyclodextrin with LiH for 24 hours and subsequently treating with one equivalent of an alkyl halide, in the aforementioned study's case, propargyl bromide. This modification in itself is of interest to our aims, because a propargyl group can be easily modified via click reaction with an azide[13–15]. To this end, we sought out to perform the modification ourselves, and proceed to possibly click reactions afterwards.

3.3.1 Synthesis of propargylated β-cyclodextrin

b-cyclodextrin and propargyl bromide (in toluene) were obtained from Sigma-Aldrich. To begin, 5 grams of β-cyclodextrin were dissolved into 85 mL of DMSO, followed by 55 mg of LiH. This was stirred under nitrogen at room temperature until the solution became clear (about 24 hours). Following this, 450 uL of a solution of propargyl bromide in toluene (80% w/w) was added alongside a catalytic amount of NaI. The mixture was stirred at 55 C in darkness for 5 hours. The solution was then poured into a liter of acetone, with the precipitate washed and filtered with acetone. The crude was filtered on a silica gel column using CH 3 CN-H 2 O-NH 4 OH (30 % v/v) 10:5:2 as an eluent to yield the final desired product as a white solid. NMR spectra were collected at 298 K on a 600 MHz Bruker Avance III spectrometer. All studies were run in water with 10% D2O. Key observed peaks in the 1H NMR spectra taken of the resultant product

were d 5.4 (H-1^{1}), 5.1 (H-1^{II-VII}), 4.45 (CH2, propargyl), 4.12 (H-3^{1}), 3.97 (H-3^{II-VII}), and 2.94 (H, terminal alkyne CH).

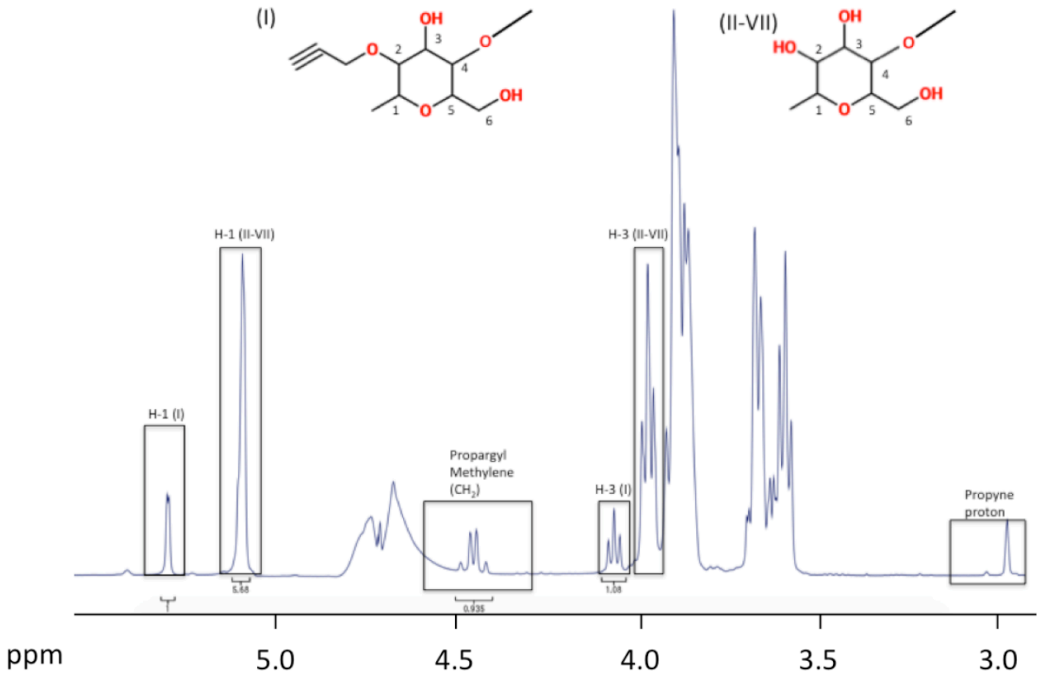


Figure 3.3: 1-H NMR spectrum of propargylated β-cyclodextrin.

Overall, the synthesis pathway was successful, and yielded around 30% product. This provided a suitable building block upon which further functionalization could be achieved. Our next step was to attempt to click a pentanoate onto the secondary face of the β-cyclodextrin, which could potentially interact with cationic adamantyl compounds previously tested with b-cyclodextrin.

3.3.2 Click reaction of propargylated β-cyclodextrin

One gram of the propargylated β-cyclodextrin was dissolved into 5 mL of water. Separately, tris(3-hydroxypropyltriazolylmethyl)amine (THPTA) was added to water to

create a stock solution at 200 mM, an CuSO4 was added to water to create a stock solution at 100 mM. The CuSO4 (5 mL) was mixed with the THPTA ligand (10 mL) and allowed to stand for several minutes. An excess of pentanoic acid was pipetted into the solution of propargyl β-cyclodextrin. The THPTA/CuSO4 solution was added (2 mL) to the solution of propargyl β-cyclodextrin. Additionally, 4 mL of 100 mM sodium ascorbate was added to the solution. The reaction was allowed to run at room temperature for an hour before being applied to a silica gel column and run using the same eluent as described above. The resultant purified compound was analyzed on a Bruke Daltonics MicrOTOF-QII spectrometer. A peak corresponding to the targeted product was identified, however, upon trying to dry down the solution containing the product, it was discovered that the solid obtained was not soluble.

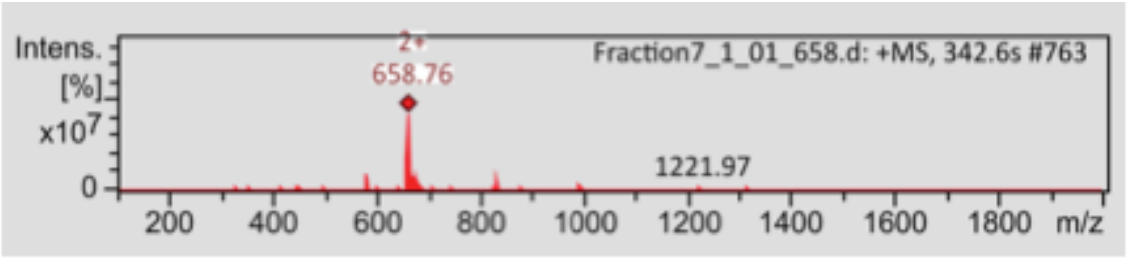


Figure 3.4: TOF Mass spectrometry of purified clicked b-cyclodextrin. The peak at 658.7 corresponds to a product of 1317 Da, indicating presence of the clicked compound in this fraction.

Overall, while the single-site functionalization with the propargyl group was successful, the click reaction yielded a product that was insoluble in water. Interestingly, the functional group that was added, a pentanoic acid, was expected to increase solubility, not decrease it, since it should deprotonate at neutral pH. We speculated that the product possibly crystallized in such a way to form highly stable isomorphs, as the

pentanoic acid arm of the product may have come to rest in its own cyclodextrin's cavity, or a neighboring cavity.

Additionally, we have been able to create a mono-3-carboxypropionamido-β-cyclodextrin (CP-β-CD), another mono-functionalized b-cyclodextrin variety. This host was synthesized by reacting succinic anhydride with 3-amino-b-cyclodextrin, a convenient precursor molecule available for purchase. The details of this derivative can be found in a report which also details binding to rimantadine[16]; the dissertation was co-investigator and second author on this publication.

3.4 REFERENCES

1. Cram, D. J. & Cram, J. M. Host-Guest Chemistry: Complexes between organic compounds simulate the substrate selectivity of enzymes. *Science* **183**, 803–809 (1974).

2. Babjaková, E., Branná, P., Kuczyńska M., Rouchal, M., Prucková, Z., Dastychová, L., Víchab J., Vícha, R. An adamantane-based disubstituted binding motif with picomolar dissociation constants for cucurbit[n]urils in water and related quaternary assemblies. *RSC Adv.* **6**, 105146–105153 (2016).

3. Fenley, A. T., Henriksen, N. M., Muddana, H. S. & Gilson, M. K. Bridging Calorimetry and Simulation through Precise Calculations of Cucurbituril–Guest Binding Enthalpies. *J. Chem. Theory Comput.* **10**, 4069–4078 (2014).

4. Ling, X., Saretz, S., Xiao, L., Francescon, J. & Masson, E. Water vs. cucurbituril rim: a fierce competition for guest solvation. *Chem. Sci.* **7**, 3569–3573 (2016).

5. Liu, S., Whisenhunt-loup, S. E., Gibb, C. L. D. & Gibb, B. C. An improved synthesis of 'octa-acid' deep-cavity cavitand. *Supramol. Chem.* **23**, 480–485 (2011).

6. Rekharsky, M. V. & Inoue, Y. Complexation Thermodynamics of Cyclodextrins. *Chem. Rev.* **98**, 1875–1918 (1998).

7. Lis-Cieplak, A., Sitkowski, J. & Kolodziejski, W. Comparative Proton Nuclear Magnetic Resonance Studies of Amantadine Complexes Formed in Aqueous Solutions with Three Major Cyclodextrins. *J. Pharm. Sci.* **103**, 274–282 (2014).

8. Hooley, R. J., Van Anda, H. J. & Rebek, J. Cavitands with Revolving Doors Regulate Binding Selectivities and Rates in Water. *J. Am. Chem. Soc.* **128**, 3894–3895 (2006).

9. Purse, B. W. & Rebek, J. Functional cavitands: Chemical reactivity in structured environments. *Proc. Natl. Acad. Sci.* **102**, 10777–10782 (2005).

10. Carrazana, J., Jover, A., Meijide, F., Soto, V. H. & Vazquez Tato, J. Complexation of adamantyl compounds by beta-cyclodextrin and monoaminoderivatives. *J. Phys. Chem. B* **109**, 9719–9726 (2005).

11. Saokham, P., Muankaew, C., Jansook, P. & Loftsson, T. Solubility of Cyclodextrins and Drug/Cyclodextrin Complexes. *Molecules* **23**, 1161 (2018).

12. Casas-Solvas, J. M. & Vargas-Berenguel, A. Synthesis of a β-cyclodextrin derivative bearing an azobenzene group on the secondary face. *Tetrahedron Lett.* **49**, 6778–6780 (2008).

13. Hein, C. D., Liu, X.-M. & Wang, D. Click Chemistry, A Powerful Tool for Pharmaceutical Sciences. *Pharm. Res.* **25**, 2216–2230 (2008).

14. Presolski, S. I., Hong, V. P. & Finn, M. G. Copper-Catalyzed Azide–Alkyne Click Chemistry for Bioconjugation. *Curr. Protoc. Chem. Biol.* **3**, 153–162 (2011).

15. Rostovtsev, V. V., Green, L. G., Fokin, V. V. & Sharpless, K. B. A Stepwise Huisgen Cycloaddition Process: Copper(I)-Catalyzed Regioselective "Ligation" of Azides and Terminal Alkynes. *Angew. Chem. Int. Ed.* **41**, 2596–2599 (2002).

16. Kellett, K., Kantonen, S. A., Duggan, B. M. & Gilson, M. K. Toward Expanded Diversity of Host–Guest Interactions via Synthesis and Characterization of Cyclodextrin Derivatives. *J. Solut. Chem.* **47**, 1597–1608 (2018).

4 EVALUATION AND MINIMIZATION OF UNCERTAINTY IN ITC BINDING MEASUREMENTS: HEAT ERROR, CONCENTRATION ERROR, SATURATION, AND STOICHIOMETRY

Biochimica et Biophysica Acta 1861 (2017) 485–498

Contents lists available at ScienceDirect

Biochimica et Biophysica Acta

journal homepage: www.elsevier.com/locate/bbagen

Evaluation and Minimization of Uncertainty in ITC Binding Measurements: Heat Error, Concentration Error, Saturation, and Stoichiometry

Samuel A. Kantonen, Niel M. Henriksen, Michael K. Gilson *

Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California San Diego, 9500 Gilman Drive, La Jolla, California 92093-0736, USA

ARTICLE INFO

Article history:
Received 16 June 2016
Received in revised form 27 August 2016
Accepted 2 September 2016
Available online 15 September 2016

Keywords:
ITC
Error evaluation
Error minimization
Data analysis

ABSTRACT

Background: Isothermal titration calorimetry (ITC) is uniquely useful for characterizing binding thermodynamics, because it straightforwardly provides both the binding enthalpy and free energy. However, the precision of the results depends on the experimental setup and how thermodynamic results are obtained from the raw data.
Methods: Experiments and Monte Carlo analysis are used to study how uncertainties in injection heat and concentration propagate to binding enthalpies in various scenarios. We identify regimes in which it is preferable to fix the stoichiometry parameter, N, and evaluate the reliability of uncertainties provided by the least squares method.
Results: The noise in the injection heat is mainly proportional in character, with ~1% and ~3% uncertainty at 27C and 65C, respectively; concentration errors are ~1%. Simulations of experiments based on these uncertainties delineate how experimental design and curve fitting methods influence the uncertainty in the final results.
Conclusions: In most cases, experimental uncertainty is minimized by using more injections and by fixing N at its known value. With appropriate technique, the uncertainty in measured binding enthalpies can be kept below ~2% under many conditions, including low C values.
General Significance: We quantify uncertainties in ITC data due to heat and concentration error, and identify practices to minimize these uncertainties. The resulting guidelines are important when ITC data are used quantitatively, such as to test computer simulations of binding. Reproducibility and further study are supported by free distribution of the new software developed here.

1. Introduction

The thermodynamic characterization of noncovalent binding interactions is important in both basic and applied fields, including molecular biophysics and drug design. While binding free energies, ΔG, can be measured by a range of methods, such as spectrometry, surface plasmon resonance [1], and enzyme inhibition [2], the method of isothermal titration calorimetry (ITC) [3–8] is unique in providing not only the binding free energy but also the binding enthalpy, ΔH, from measurements taken at a single temperature. Binding enthalpies are of interest because they offer further information regarding the physics of noncovalent interactions, such as the phenomenon of entropy-enthalpy compensation [9–13], and can bear on mechanisms of protein folding and drug action. For example, modes of drug-receptor binding can sometimes be assigned based on knowledge of both ΔG and ΔH [14,15]. Moreover, the determination of ΔH, in addition to ΔG, expands the dataset available to test and improve the accuracy of molecular simulations: much as heats of vaporization and sublimation played a central role in defining the force field parameters used for liquid-state simulations [16–19], so binding enthalpy data, particularly for experimentally and computationally tractable host-guest systems, should be useful to test and improve the force fields used for simulations of noncovalent binding [20,21]. For all of these applications, it is desirable to use procedures that minimize the uncertainty in the thermodynamic results and to have a quantitative understanding of these uncertainties.

Sources of uncertainty in ITC data may include [3,22–28] errors in the quantification of released heat by the instrument; errors in the concentrations of the solutions; baseline error, due to instrument problems or environmental perturbations; nonideality of the solutions; and procedural problems, such as the presence of air bubbles in the cell or use of non-optimal stirring speeds. Each of these sources of error can produce its own pattern in the data, and can propagate in its own way to the fitted binding free energy and enthalpy. For example, procedural and baseline errors can yield abnormal-looking enthalpograms, whereas noise in the measured

Abbreviations: ITC, isothermal titration calorimetry; LS, least squares; MC, Monte Carlo.

* Corresponding author.
E-mail address: mgilson@ucsd.edu (M.K. Gilson).

http://dx.doi.org/10.1016/j.bbagen.2016.09.002
0304-4165/

heat released from successive injections can produce a jagged or noisy-looking Wiseman plot. Concentration error, in contrast, does not yield obviously pathological graphs, but can nonetheless propagate to the fitted thermodynamic results. Finally, as already widely recognized [7,29,30], when molecules release or take up protons on binding, the consequent uptake or release of protons by the buffer can lead to added enthalpic changes which obscure the heat signature of the binding event itself. The magnitude of this effect depends on the buffer; for example, it is particularly large for Tris [31], and small for phosphate [31]. Arguably, buffer ionization heats do not lead to experimental errors, but rather to errors of interpretation.

The literature affords divergent views of the level of uncertainty in ITC data. On one hand, it is not uncommon to see published ITC data with reported experimental uncertainties in the hundredths of kcal/mol. On the other hand, the ABRF-MIRG'02 study [32], in which 17 laboratories independently measured the binding thermodynamics of the same protein and small molecule, suggested that experimental uncertainties might be as high as 20%. Much of the error in this study appears to be procedural, as omitting the four results for which the representative enthalpograms and/or Wiseman plots are obviously abnormal or noisy (graphs 2, 7, 8 and 14 in Fig. 4 of reference 32) leads to a less problematic standard deviation of 11% in ΔH. Further analysis led to the suggestion that that errors in the concentration of the syringe reactant represented a major source of the variation in the reported binding enthalpies [25]; and the potential sensitivity of binding enthalpy results to errors in the syringe concentration has been emphasized elsewhere as well [26,27]. Importantly, the errors observed in the ABRF-MIRG'02 study are not necessarily representative of typical published ITC data, as the participants were not required to be experts, and novices were explicitly welcomed and included (Michael Doyle, personal communication).

The uncertainties in binding free energies and enthalpies obtained from ITC experiments can be mitigated by appropriate experimental design, execution, and analysis [7,22–24,26,33,34]. Thus, as early noted [3], it is helpful to design experiments with a C value – the binding constant times the cell concentration and stoichiometry – within an appropriate range, though it has also been argued that experimental data with low C values can still yield useful thermodynamic data [34]. Additionally, good laboratory practices can minimize uncertainties stemming from environmental noise; one should be cautious in applying ITC when the heat of reaction is small in magnitude, as the small magnitudes of the injection heats being measured in such cases can lead to a low signal-to-noise ratio; and pathological enthalpograms should be discarded. Note, too, that sophisticated methods of baseline determination have been developed to minimize error that may stem from drift or inconsistencies across experiments [24].

The method used to fit theoretical curves to the experimental data also can affect uncertainties in ITC results. For example, fixing the binding stoichiometry, N, in the data analysis process, rather than allowing it to be optimized as an additional fitting parameter, may reduce experimental error for experiments at low C values [34]. On the other hand, allowing N to float, instead of fixing it, is widely regarded as a useful means of reducing the sensitivity of the results to errors in the concentration of the cell solution [25]; accordingly, the software distributed with many ITC instruments defaults to this option. (Another approach is to use global fitting over multiple Wiseman plots, obtained under different experimental conditions, keeping N fixed at a known integral value and treating concentration error explicitly [35,36].) The problem of determining whether to treat N as fixed or let it float is complicated by the fact that the commonly used least squares (LS) error analysis method does not propagate concentration error to the reported thermodynamic results, as recently highlighted [26], so it might provide misleading guidance as to which curve-fitting procedure is best. It is therefore of interest to use more robust error propagation methods, such as Monte Carlo simulations [22,33,37] to study how syringe and cell concentration error affect errors in the binding free energy and enthalpy, and how this propagation is influenced by the treatment of N in the curve fitting process.

Another source of error in ITC measurements is heat error, the noise in the heat measured for each injection. This can result from random variations in the injection volumes, relative to what was intended; or from noise in the measurement of the heat released from each injection. Like concentration error, heat error propagates through to the thermodynamic results, and it has been argued that using fewer ITC injections can mitigate the resulting errors [22]. The analysis was predicated on a specific model of heat error, with a fixed 0.28 μcal/injection measurement error and an additional contribution due to error in the injection volume, but the literature contains additional models for heat error. For example, prior studies have assumed fixed errors of 0.6 μcal/injection [33], 0.25 μcal/injection [7] and 0.1 μcal/injection [37]; another assumed a mixed error model, in which injections with smaller heat release were associated with a fixed error of 0.5 μcal/injection, but injections with heat release >100 μcal were associated with a proportional error of 0.5% [34]. Another study included a proportional term in the overall estimate of error, but it was insignificant for heat releases smaller than 300 μcal [38]. Wiseman and coworkers examined the variance in measured heat for a series of small electrical pulses of fixed magnitude [3], but did not report how the variance changed as a function of the amount of heat released. Additionally, the heat error reported for a series of water-into-water injection was found to be about 0.06 μcal, considerably less than predicted by some of the assumed fixed errors listed above [39]. Thus, it is not clear that the literature contains empirical data that would direct one toward any particular model for the noise in the heat measured per injection.

The present study addresses a number of these issues regarding the design and analysis of ITC experiments, focusing on a model case of 1:1 binding. In particular, we empirically evaluate uncertainties in injection heats and solution concentrations, and use Monte Carlo data modeling to examine how these uncertainties propagate to uncertainties in the derived binding enthalpies, for experiments at low and optimal C values, and with the value of N fixed or allowed to float. We focus primarily on binding enthalpies as opposed to binding free energies, because (as shown below) free energies are comparatively insensitive to experimental error, while errors in binding enthalpies are generally larger and present more of a challenge to fit with high precision [26,40]. However, we also provide results showing uncertainty in binding free energies given realistic experimental errors, as part of a global examination of errors encountered in what might be a typical ITC experiment. The present results have implications for how uncertainties in binding thermodynamics are assigned, and for how to design and analyze ITC experiments in a manner that reduces these uncertainties.

2. Materials and Methods

2.1. Materials

β-cyclodextrin (catalog no. C-4767, purity >97%) was obtained from Sigma-Aldrich Company (St. Louis, MO). Initial batches of β-cyclodextrin that we had stored for over a year showed some evidence of aggregation [41] in newly made concentrated solutions that had stood for on the order of an hour or more, based on slight clouding and confirmatory spectrophotometry, so a new batch was purchased and used for the present experiments. Solutions made from the new lot showed no evidence of aggregation. Thermogravimetric analysis of the initial batch of β-cyclodextrin confirmed Sigma-Aldrich's specification of 8 mol of water per mole of β-cyclodextrin, and the amounts weighed to achieve a target concentration were adjusted to compensate for the water content. NMR spectroscopy was used to confirm the structure and purity of the new lot of β-cyclodextrin. Succinic acid and NaOH were both obtained from Fisher Scientific through the chemistry stockroom at University of California, San Diego.

2.2. Isothermal Titration Calorimetry Experiments

ITC experiments were performed using a MicroCal model VP-ITC (MicroCal, Northampton, CT, Serial Number 01–08-930). It has been emphasized that even seemingly minor deflections in the power baseline during an ITC experiment can lead to nontrivial errors whose remediation, when feasible, requires detailed analysis [24]. We therefore sought to minimize environmental sources of baseline noise. We found that small, deliberate vibrations of the laboratory bench on which the calorimeter was set caused significant deflections in baseline. To prevent this, the calorimeter was placed in an isolated room on a 2″ thick block of urethane foam. This setup eliminated deflections during deliberate vibration of the bench and provided a more stable baseline. A purpose-built, clear acrylic shield was furthermore used to reduce possible temperature shifts due to drafts. The instrument's built-in Y-axis calibrations were performed, at 27 °C, to verify that the instrument was responding to known power inputs within the 1% error tolerance prescribed by the manufacturer.

Following a previously published procedure [42], we determined the uncertainty of the heat release per injection by repeatedly injecting fixed amounts of succinic acid solution into an excess of NaOH and determining the variance of the integrated heats across the series of injections. In principle, all injections should yield essentially the same heat, due to the excess of NaOH. The succinic acid and NaOH solutions were prepared at 2 mM and 20 mM respectively. Masses of dry succinic acid were measured with a Sartorius CPA225D Micro Balance, while the NaOH solution was prepared by dilution of 1 M stocks as purchased. Solutions were prepared in 25 mL volumetric flasks. Measurements were taken with injection volumes of 10, 5, and 2.5 μL injections, with twenty injections per experiment. The reference power was set at 25 μcal/sec for all experiments. In order to determine whether the observed drop in variance with diminishing injection volume in these initial experiments should be attributed to variations in injection volume versus variations in the measurement of released heat, additional experiments were done in which the injection volumes were held at 10 uL, and the concentration of succinic acid was instead reduced to generate solutions ranging from 20 mM to 0.5 mM, leading to reduced heat release per injection. We additionally performed water-into-water experiments, which generate very small injection heats, to look at the minimal variance in injection heat, at temperatures of 27 °C and 65 °C. Origin 7.0 was used to integrate the heat released from the injection heats. For 2.5 μL injections, the first two injections were discarded; for 5 and 10 μL injections, the first injection was discarded. This is due to the typically smaller first injection in ITC experiments, possibly due to diffusion of syringe material into the cell during equilibration.

2.3. Measurement of Concentration Uncertainty

The solutions of β-cyclodextrin for concentration error analyses were prepared in 250 mL volumetric flasks, with 2.00 g β-cyclodextrin per 250 mL batch of solution; and in 25 mL volumetric flasks, with 0.200 g of β-cyclodextrin per 25 mL batch. Masses were measured with a Sartorius CPA225D Micro Balance. Each solution of β-cyclodextrin was prepared from the same stock of 0.5% (v/v) acetonitrile in water (prepared prior to dissolution of weighed β-cyclodextrin at 1 L), to provide a consistent standard peak for NMR measurements. Proton NMR measurements were carried out on a Varian NPA600, using 10% deuterium oxide in the 0.5% acetonitrile solution used to make up each sample of β-cylodextrin, in order to lock to ^{2}H. To evaluate the uncertainty in the concentrations of the cyclodextrin solutions, we prepared four independent solutions at each volume (250 mL, 25 mL) as described above, obtained the proton NMR spectrum of each solution, and integrated the five peaks corresponding to glucose protons, using Bruker TopSpin software, as well as the acetonitrile peak. The ratio of the peak integrals from the β-cyclodextrin to the peak integral of acetonitrile was measured, and compared amongst solutions to determine the overall standard deviation of the concentration of solutions. This does not provide systematic error in concentrations, but rather the variation of concentrations amongst four solutions prepared using identical methods. Note that spectrophotometry could not be used for this purpose, because β-cyclodextrin does not absorb significantly in the UV–Vis range.

2.4. Generation and Analysis of Artificial ITC Data

We generated and analyzed artificial Wiseman plots based on thermodynamic parameters for the reversible association of a model system previously studied by both our laboratory and others [43]: the adamantyl-based drug rimantadine with the cyclic oligosaccharide β-cyclodextrin, with thermodynamic binding parameters $\Delta H = -6500$ cal/mol and $K = 36{,}000\ M^{-1}$, $N = 1$. The data were modeled for our own (unpublished) experimental setups with rimantadine in the cell and cyclodextrin in the syringe, with a favorable C value of 50 ($C_{syringe} = 13$ mM, $C_{cell} = 1.4$ mM) and an unfavorable C value lowered to 0.50 by lowering concentrations ($C_{syringe} = 3.5$ mM, $C_{cell} = 0.014$ mM). Except as otherwise noted, the simulated experiments used 25 injections of 10 μL each, leading to maximum saturation of 99% for both the low and high C cases. In selected cases, we also modeled low C experiments run to only 55% saturation ($C_{syringe} = 0.32$ mM, $C_{cell} = 0.014$ mM), where the amount of injectant was diminished by reducing the concentration in the syringe. No heat of dilution was incorporated into the simulated Wiseman plots, so the nominal injection heats provided by the model derive only from the binding reaction.

The equations for a single set of identical sites [3,44] were implemented in Python scripts, which produced ideal Wiseman plots based on the desired K, ΔH, concentration values, injection volumes and numbers of injections. The volume of the ITC cell was set to that of our VP-ITC instrument, 1.4614 mL, and the expressions used accounted for dilution of the sample cell during the experiment, as typically done by Origin 7.0. We verified that fitting artificial plots based on these parameters, with both Origin 7.0 and our Python code, returned the thermodynamic parameters used to generate them. The resulting Wiseman plots were analyzed by additional Python scripts that automated the subsequent curve-fitting procedures, using an implementation of the Marquardt–Levenberg nonlinear optimization method available in the SciPy library. This procedure yields best-fit values of K and ΔH for a given Wiseman plot, along with LS uncertainties like those provided by the Origin software.

2.5. Error Analysis by Monte Carlo Sampling

Experimental uncertainties in ITC results are commonly estimated from the residuals and partial derivatives of the LS fit. For example, it is the LS uncertainties that are reported by the widely used Origin software [44]. However, as detailed in Results, this method is itself subject to error, tending to yield uncertainties that are unrealistically low. Here, therefore, we follow others [7,22,33] in using a more rigorous Monte Carlo sampling method to determine how uncertainties in concentration and measured heat release propagate to the reported thermodynamics quantities. These reference results are then compared with the more commonly reported LS uncertainties to assess the limitations of the latter. The Monte Carlo method was implemented as follows. The uncertainty in the injection heats was modeled by associating the heat release for each injection with a Gaussian distribution of heat release with the ideal mean value, μ_Q, and a user-defined standard deviation σ_Q following Eq. (1) (below). Two thousand different plots of raw heats were generated by randomly resampling the raw heat value at each point from its respective Gaussian. We confirmed for several test cases that increasing the number of plots from 2000 to 10,000 led only to ~0.1% changes in reported binding enthalpy uncertainties,

for even the most error prone conditions. Concentration error was modeled by using syringe and cell concentrations drawn from Gaussian distributions, with their means set to their nominal concentrations, μ_C and their standard deviations set to a user-defined value, σ_C. In each case, the resampled raw heats were normalized based on the erroneous syringe concentration to generate a new Wiseman plot. Values of K and ΔH were then fitted to each resulting Wiseman plot, using the Python script described above, to yield a statistical distribution of these fitted quantities, from which we computed the mean and standard deviation of ΔH (μ_H and σ_H, respectively) and ΔG (μ_G and σ_G), where the spreads are determined by the uncertainties in heat release and solution concentration, as manifest in the resampled Wiseman plots. The LS fitting routine rarely generated pathological results, with very high reported LS errors, apparently due to occasional failure of the solver. To avoid incorporating these pathological cases into the statistics, results for which the reported LS error was in the top first percentile were discarded.

3. Results

This section first reports on the experimental characterization of heat error, and how this error, in the absence of concentration error, propagates to measured binding thermodynamics, for both normal (C value = 50) and low (C value = 0.5) C value experiments. It then reports, analogously, on measurements of concentration error and on how concentration error without heat error propagates to the binding enthalpy. Finally, it reports on global analyses in which both heat and concentration error are present, in order to gain insight into how to assess and minimize uncertainty in measurements of binding thermodynamics. Although we focus primarily on how errors propagate to binding enthalpies, as these are more susceptible to error than free energies, selected analyses are also done for the free energy, and Section 3.4 assesses errors in both enthalpy and free energy for a range of realistic experimental conditions.

3.1. Heat noise in enthalpograms: character and propagation

Random noise in the integrated peaks of the VP-ITC enthalpogram was measured in terms of the peak-to-peak variation across a series of nominally uniform peaks, which were generated by injection of 2 mM aqueous succinic acid into 20 mM sodium hydroxide solution. Due to the excess of hydroxide in the cell, each injection of succinic acid should be equally exothermic, with only slight loss of heat over the course of a single experiment due to dilution of the hydroxide, as evident in sample data (Fig. 1a); see SI for details (Supplementary Table 1). We find that the standard deviation of the injection heats, σ_Q, across a series of nominally identical injections, at 27 °C, is approximately proportional to the mean heat released during the injection, Q, across injections of various volumes (Fig. 1b, black circles). The slope of a linear fit to these data is a temperature dependent coefficient of proportional error, with the form $\zeta(27) = 0.01$, which corresponds to 1% uncertainty in the heat measurements. Note that the range of heats in the succinic acid/NaOH experiment (Fig. 1b) is in the middle of that occurring in the ITC experiments at optimal C modeled in this study. Because we have observed noisier enthalpograms at higher temperatures, we also examined the standard deviation of the peaks at 65 °C. As shown in Fig. 1b (red circles), the error, about 3% of the measured heat, is significantly greater than at 27 °C; we estimate $\zeta(65^{\circ}) \approx 0.03$, based on the locations of the red "x" and the red circles in Fig. 1b.

Additionally, water-into-water experiments were performed to assess any underlying error, since water-into-water injections generate extremely small amounts of heat (Fig. 1b, c, x's). This minimum level of noise presumably is associated with fluctuations in the calorimeter's baseline signal, perhaps coupled to variations in measured Joule heating. As this irreducible heat variance potentially results from processes different from those that generate the proportional noise observed in the present succinic acid/NaOH study, we assume that both processes operate simultaneously and independently. As a consequence, the total variance of the heat per injection should be estimated as the sum of the variances of the two processes, so that their standard deviations add in quadrature:

$$\sigma_Q(T) \approx \left\{ [\zeta(T)Q]^2 + \sigma_{irr}(T)^2 \right\}^{\frac{1}{2}} \quad (1)$$

Here, σ_Q is the standard deviation of the measured injection heat; $\zeta(T)$ is the temperature dependent coefficient of the proportional error component; and $\sigma_{irr}(T)$ is the standard deviation of the temperature-dependent irreducible heat noise. Adding the two variances affords an error model which is conservative, relative to prior models that assign the irreducible noise minimum as the floor of a proportional model [34,37] rather than as an additive noise component. Our water-into-water studies yield $\sigma_{irr}(27)$ of 0.13 μcal, and $\sigma_{irr}(65)$ of 0.39 μcal, with a mean heat release of 0.36 and 0.70 μcal, respectively. These may be compared with a prior report of 0.06 μcal [39] for $\sigma_{irr}(27)$.

Because the differences in mean heat reported in the prior paragraph and illustrated in Fig. 1b were generated by changing the injection volume while using a single syringe solution, it was possible that the increase in noise with increased heat release resulted not from the

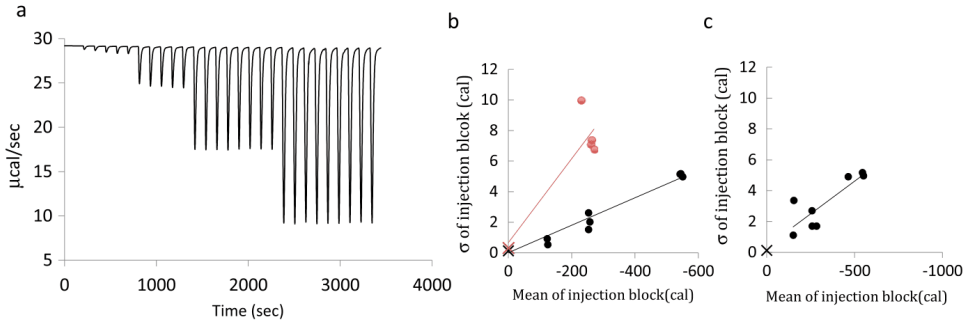


Fig. 1. Injection of succinic acid into NaOH. (A) A raw enthalpogram of varying injection sizes of 2 mM succinic acid into 20 mM NaOH. Peaks shown (from smallest to largest) are 0.5 μL, 2.0 μL, 5 μL, and 10 μL. The slight increase in peak size on the 0.5 and 2.0 μL peaks stems from the well-known diffusion of syringe reactant prior to the start of the experiment. As noted in the Methods section, for experimental data the first one or two injections were discarded. (B) Plots of means of heat vs. standard deviations with varying volume. Black points show data at 27 °C, red points at 65 °C. The black and red Xs are water-water experiments at 27 °C and 65 °C, respectively. (C) Plots of means of heats vs. standard deviation for constant volume succinic acid injections. The black X is a water–water experiment at 27 °C.

increase in heat release as such, but instead from the increase in the volume of fluid injected, due to some property of the syringe mechanism. We therefore checked whether the uncertainty in the enthalpogram would remain proportional to peak size if we kept the injection volume constant (10 μL) and instead adjusted the heat per injection by using solutions of different concentrations. We found that, across injections of constant volume, the standard deviation of the heat per injection, σ_Q, still is approximately proportional to the mean heat release, μ_Q, with a constant of proportionality $\zeta(T) = 0.008 - 0.009$ (Fig. 1c), depending on whether or not the linear fit was forced through the origin. Because these values are close to the prior value of 0.01 (above), we conclude that the increase in noise with increasing heat release per injection is attributable to noise in the heat measurement rather than to changes in the injection volume. This result appears to disagree with a prior model in which a non-fixed component of the heat error was considered to result from variance in the injection volumes [22].

It is of interest to determine whether the heat measurements from successive injections were mutually correlated, because such correlation could affect the estimated error of the fitted thermodynamic quantities [22]. For three independent series of 10 μL injections, the R^2 values between successive (i, i + 1) injection heats were 0.05, 0.33, and 0.07, and, for three independent series of 5 μl injections, the analogous R^2 values were 0.50, 0.27, and 0.39. Substantially smaller values of R^2 were observed for the ith and (i + 2)th heats. The fitted slopes of linear regressions between the ith and (i + 1)th heats were positive, with values on the order of 0.5. Thus, there appears to be a weak correlation between successive injection heats, but the correlation coefficients are variable from one study to the next, the correlation does not persist across multiple peaks, and the degree of correlation appears to be greater for smaller peaks. Given the modest and inconsistent strengths of the correlations, we treated successive peaks as uncorrelated when modeling experimental error in the present study. Sample correlation plots are shown in the SI (Supplementary Fig. 1).

We next used Monte Carlo sampling to study how the heat errors reported above propagate to uncertainties in the derived values of ΔH. As detailed in Methods, multiple artificial Wiseman plots were generated, each with its own set of random, uncorrelated Gaussian errors added to the injection heats, using the proportional error model of Eq. (1), for values of ζ ranging from 0 to 0.03, and for values of C equal to 50 and 0.50. As mentioned above, an irreducible error was added in quadrature to the proportional component. Based on the previously mentioned water–water experiments, the baseline error was set as $\sigma_{irr} = 13\zeta\ \mu cal$. Samples of these artificial Wiseman plots (with $\zeta = 0.01$) are provided in the SI (Supplementary Fig. 2). A separate least squares (LS) fitting for each Wiseman plot, with N either fixed or allowed to float, yielded a value of ΔH and a report of uncertainty from the LS fit. The mean, μ_H, and standard deviation, σ_H, of the resulting values of ΔH, were computed, and we also computed the mean of the uncertainties reported by the LS method across each set of MC samples. With the present model for heat error (Eq. (1)), we observed a small exothermic bias in μ_H relative to its nominal value, rising to a maximum of about 7% at $\zeta = 0.03$, in cases where N is allowed to float and C = 0.5 (Supplementary Fig. 4). However, the scatter of the fitted enthalpies, σ_H, in this case is so large that it effectively drowns out the bias (Supplementary Fig. 4). This bias disappears at higher C values and when N is fixed.

As shown in Fig. 2, the uncertainty in the enthalpy, σ_H, depends strongly on the values of ζ and C, and on whether N is treated as fixed or floating. Thus, when C has the favorable value of 50 and N is allowed to float, as is commonly done, both the Monte Carlo and the LS estimates of σ_H remain below 2%, even when the proportional part of the heat error rises to 3% ($\zeta = 0.03$), the value observed at 65 °C; see solid black and red lines, respectively, in Fig. 2a. However, when C has the unfavorable value of 0.5, and N is allowed to float, both the Monte Carlo and LS estimates of σ_H are much higher. Indeed, the Monte Carlo result goes above 20% when the heat error is 1.5% (dashed black line, Fig. 2a). Additionally, LS uncertainties underestimate the reference Monte Carlo results, as they reach only about half the respective Monte Carlo error estimate. Fixing the value of N at its nominal value of 1 somewhat reduces the already low errors when C = 50 (black and red solid lines, Fig. 2b), and markedly reduces the uncertainty in the enthalpy when C = 0.5 (black and red dashed lines, Fig. 2b). The observation that fixing N at its nominal value can reduce the propagation of heat error is consistent with prior observations [34]. The results for C = 0.5 in Fig. 2 were obtained in modeled experiments where the cell reactant was 99% saturated with the syringe reactant at the end of the titration. Results for experiments reaching only 55% saturation are furthermore provided in Tables 2 and 3.

3.2. Concentration error: characterization and propagation

We empirically evaluated the uncertainty in the concentrations of stock solutions of β-cyclodextrin made up in 250 mL and 25 mL volumetric flasks. In each case, we made four solutions of β-cyclodextrin and took NMR spectra, integrated peaks corresponding to glucose protons and the reference acetonitrile peak, and computed the standard deviations of the ratio of the glucose peaks to the reference peak, as described in Methods. This procedure yielded a relative uncertainty, $\sigma_C/Conc$, of 0.6% in the concentration, when solutions are prepared in

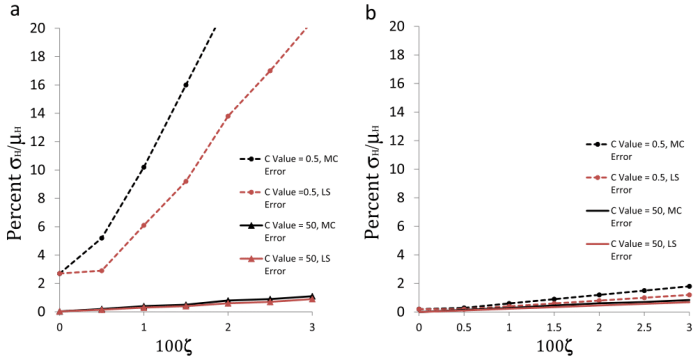


Fig. 2. Relative uncertainty in ΔH as a function of heat error coefficient, ζ, with $\sigma_{irr} = 13\zeta$ μcal added in quadrature. Black lines are uncertainties from Monte Carlo resampling; red lines are mean uncertainties provided by LS fitting only. Solid lines: C = 50; dashed lines: C = 0.5. a) results for floating N; b) results for N fixed at one. No concentration error is included in this analysis.

250 mL volumetric flasks, and 1.1% when solutions are prepared in 25 mL volumetrics. We then used model calculations to examine how concentration errors propagate to the reported enthalpy, in the absence of heat error, as detailed below.

Whereas heat noise leads to rough-looking Wiseman plots (Supplementary Fig. 2), concentration error yields Wiseman plots that are distorted, but smooth, as illustrated in Fig. 3, which uses syringe and cell concentration errors of ±10% to help with visualization. The curves are smooth because errors in the solution concentrations affect the entire Wiseman plot in a consistent manner. Because the x-axis of the Wiseman plot is the molar ratio of syringe to cell concentration, error in the cell concentration rescales the plot along the x-axis (Fig. 3a). Error in the syringe concentration also rescales the graph along the x-axis, though in the opposite sense from error in the cell concentration, and it additionally rescales the y-axis, due to renormalization of the heat released per mole of injectant (Fig. 3b). The effects of these ±10% concentration rescalings on the fitted values of the binding enthalpy, ΔH_{fit}, are summarized in Table 1, which includes results for not only syringe-only and cell-only concentration errors, but also their combinations. (The combination Wiseman plots are provided in Supplementary Fig. 5.) The errors are given as $\Delta H_{fit} - \Delta H_{nominal}$, where $\Delta H_{nominal}$ is the nominal binding enthalpy of −6.5 kcal/mol. A number of interesting observations emerge.

First, allowing N to float prevents propagation of cell concentration error to the binding enthalpy, but allows syringe error to propagate nearly linearly, in the sense that 10% concentration error generates about 10% enthalpy error. (The modest degree of nonlinearity presumably results from the rescaling of the Wiseman plot along not only the y-axis, but also the x-axis; see above.) Nonetheless, the LS error is always zero when N is allowed to float, whether or not the enthalpy is correct, because including N as a fitting parameter allows the smooth Wiseman plots that result from pure concentration error to be perfectly fit by theoretical curves. This is illustrated in Fig. 3, which shows how Wiseman plots shift with various concentration errors. For example, increasing the concentration in the cell to 10% above its nominal value pushes the curve to the left, because the x coordinate, X/M, becomes smaller due to the larger cell concentration. Furthermore, in cases where the syringe and cell concentration errors are equivalent (e.g., both +10%), the fitted value of N reverts to its nominal value of 1, but the actual errors in the binding enthalpy are among the largest listed in the table. The appearance of LS errors of zero for cases where the actual error is nontrivial highlights an important limitation of this statistic.

Table 1
Analysis of unsigned errors in fitted binding enthalpy (ΔH Error) due to error in syringe and cell concentration for both N floating and N fixed. Results shown are from a C value of 50, with 99% saturation. The unsigned uncertainty reported from the least squares fit (LS Err) is also shown, along with the reported value of N, when it is allowed to float. The nominal ΔH is −6.5 kcal/mol for all conditions.

Percent Conc Error		N floating				N = 1		
Syringe	Cell	ΔH_{fit}	ΔH Error	LS Err	N	ΔH_{fit}	ΔH Error	LS Err
0	0	−6.5	0.0	0.0	1.0	−6.5	0.0	0.0
−10	0	−7.2	0.7	0.0	1.1	−6.8	0.3	0.17
+10	0	−5.9	0.6	0.0	0.9	−6.4	0.1	0.14
0	−10	−6.5	0.0	0.0	0.9	−6.2	0.3	0.14
0	+10	−6.5	0.0	0.0	1.1	−7.1	0.6	0.17
+10	−10	−5.9	0.6	0.0	1.0	−5.9	0.6	0.0
+10	+10	−5.9	0.6	0.0	1.2	−7.2	0.7	0.35
−10	−10	−7.2	0.7	0.0	0.8	−6.6	0.1	0.32
−10	+10	−7.2	0.7	0.0	1.0	−7.2	0.7	0.0

When N is fixed at its nominal value of 1, both cell and syringe error can propagate to the enthalpy (Table 1), and the LS error is nonzero in most cases, because theoretical curves with $N = 1$ are unable to exactly fit experimental curves with concentration error (Fig. 3). The exception occurs when the cell and syringe errors happen to be the same; i.e., +10% or −10%. In these cases, a theoretical curve can perfectly fit the Wiseman plot, but the fitted enthalpy is wrong by about 10%. Interestingly, although fixing N = 1 yields larger LS errors than allowing N to float, the MC analysis shows that fixing N yields enthalpies that are, on average, more accurate. Thus, for the present examples, the mean unsigned errors for all of the concentration error cases in Table 1 are 0.41 for N floating and 0.38 for $N = 1$. In summary, for these model calculations, allowing N to float tends to increase actual error but decreases the error reported by the LS method. Additionally, even when N = 1, the LS error does not correlate with the actual error in the enthalpy, as shown in Fig. 4.

Finally, Fig. 5 examines how errors in the syringe and cell concentrations map to errors in the enthalpy, under conditions of optimal and low C; with N fixed or floating; and, for the low C case, experiments run to 99%, 55% and 25% saturation of the cell reactant. Several interesting points emerge from this analysis. First, as well recognized, when N is

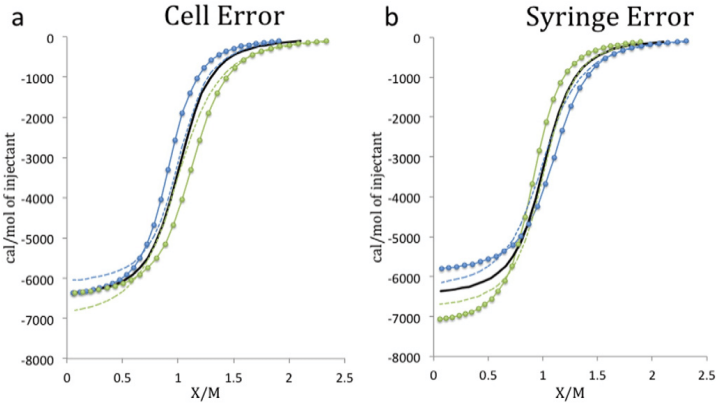


Fig. 3. Influence of concentration error on Wiseman plots (C value = 50). Black line: nominal result. Points represent Wiseman plots with 10% concentration error either added (blue) or subtracted (green) from the nominal syringe or cell concentration. Lines of the respective colors are the corresponding LS fits. Solid: N floating; dashed: N fixed at 1.

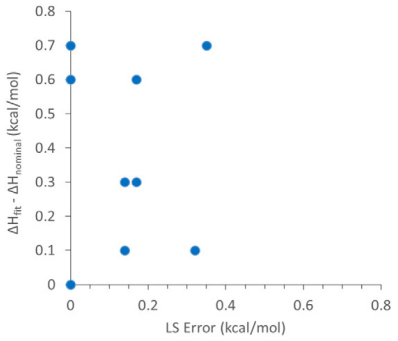


Fig. 4. Plot of actual unsigned error, abs($\Delta H_{fit} - \Delta H_{nominal}$) vs. reported LS error with fixed N; data are from Table 1.

allowed to float (Fig. 5a,c), the enthalpy remains independent of cell concentration error and varies nearly in proportion to syringe concentration error, for both optimal and low C values. For experiments at optimal C values, fixing N (Fig. 5b) increases the sensitivity of the enthalpy to the cell concentration error, while reducing its sensitivity to the syringe concentration. In addition, both dependencies have more curvature. Finally, for low C value experiments with fixed N (Fig. 5d), the sensitivity of the enthalpy to both the syringe and cell concentration errors depends on saturation, with decreasing saturation leading to increased sensitivity. In particular, whereas it was previously shown that the enthalpy can be virtually insensitive to syringe concentration at low C values and with N fixed [34], the present extended analysis shows that this holds only at high saturation, and that the sensitivity of the enthalpy to both the syringe and the cell concentration becomes more problematic for experiments run to lower saturation: the sensitivity increases slightly for a maximum saturation of 55% and increases markedly at 25% saturation.

3.3. Propagation of Combined Heat Error and Concentration Error

The prior sections examine the character and propagation of heat noise and concentration error separately. Here we use Monte Carlo sampling of heat and concentration error simultaneously to examine when or whether it is advantageous to use fewer *versus* more injections, and when or whether N should be allowed to float *versus* being treated as fixed.

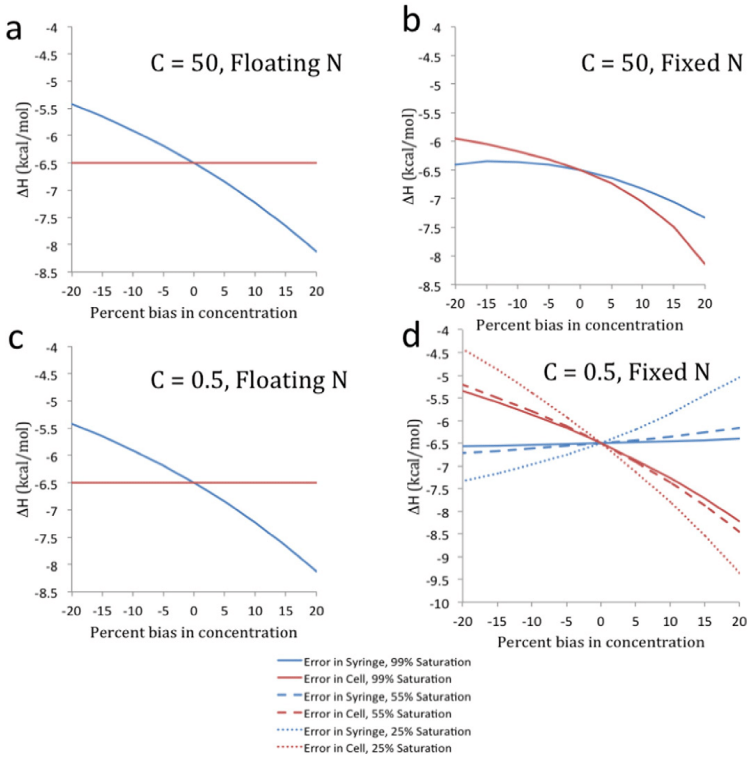


Fig. 5. Plots of fitted enthalpy with either cell or syringe concentration error applied to either Wiseman plots with C value of 50 (with N floating in a, or N fixed in b) or 0.5 (with N floating in c, or N fixed in d). For C values of 50, 99% saturation was used only. For C values of 0.5, multiple saturations (25% dotted line, 55% dashed line, 99% solid line) were used. These are only evident in panel d; all saturations gave identical results in panel c (floating N).

3.3.1. Number of injections

It has been suggested that using fewer injections reduces the propagation of heat error to the thermodynamic data [23,26]. However, the prior study was based on a different model of heat error and used an artificially large fixed component of the heat error. We therefore investigated how the number of injections affects the uncertainty σ_H when the present empirical heat error model is applied, by realistically modeling experiments that spread the same amount of injected cyclodextrin across from 5 to 40 injections, with other experimental parameters held constant. Monte Carlo sampling was used to generate artificial Wiseman plots with 1% proportional heat error ($\zeta = 0.01$), 0.13 μcal baseline error (σ_{irr}), and 0.6% concentration error, and LS curve-fitting was used to obtain the corresponding values of the binding enthalpy. Results were obtained with both optimal (C = 50) and low (C = 0.5) values of the C parameter, and the results were analyzed with N allowed to float and N fixed at its nominal value of 1. For all data points, the total heat release was kept constant.

Under all conditions examined, the uncertainty in ΔH is found to decrease when more injections are used (Fig. 6), except for a minor rise (<1%) on going from 15 to 40 injections for C = 0.5 with N fixed (Fig. 6c). The trend is strong in the case of low C when N is allowed to float (Fig. 6a), as the percent error in the binding enthalpy falls from 35 to 10% as the number of injections is increased from 5 to 40. In other cases the effect is weaker, with uncertainties improving by several tenths percent on going from 5 to 40 injections. A corresponding analysis for ΔG (Supplementary Fig. 3) reveals trends similar to those for ΔH when C is optimal (C = 50), but shows that, when N is fixed at low C values, the error in ΔG is insensitive to injection number. Overall, the combination of these two analyses indicates that increasing the number of injections is likely to improve uncertainty in both ΔG and ΔH.

3.3.2. Floating versus fixing the stoichiometry, N

The analyses above show that fixing N can reduce the propagation of heat error and syringe concentration error to the binding enthalpy, relative to letting N float, but can, on the other hand, increase the propagation of cell concentration error. As a consequence, it is not immediately clear whether or when it is best to let N float, as is commonly done, or instead to treat N as fixed, and it would be useful to understand how fixing N affects the uncertainties in fitted parameters that arise due to both heat error and concentration error. Thus, we studied this choice by using Monte Carlo sampling to evaluate the uncertainty in the enthalpy when N is allowed to float, $\sigma_{H,float}$, and when N is fixed at 1, $\sigma_{H,fixed}$, for multiple combinations of heat and concentration error, where both the syringe and cell reactants were assumed to have the same concentration error. The percentage change in uncertainty, $100(\frac{\sigma_{H,float}}{\mu_{H,float}} - \frac{\sigma_{H,fixed}}{\mu_{H,fixed}})$, is negative under conditions where floating N yields a more precise value of ΔH than fixing N, and *vice versa*. This quantity is mapped as a function of ζ(T) and percent concentration error, for low and optimal C values (Fig. 7a,b). The contour at which $\sigma_{H,float} - \sigma_{H,fixed} = 0$ divides two regions of the map: one where it is better to let N float, the other where it is better to fix N = 1.

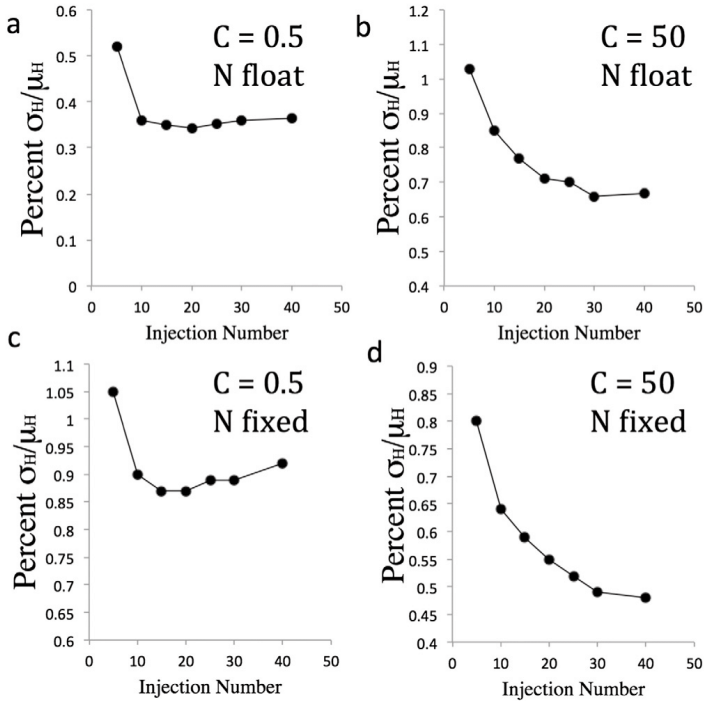


Fig. 6. Uncertainty of fitted enthalpy as a function of number of injections in model experiments. The ordinates report standard deviations of ΔH relative to the magnitude of the nominal enthalpy across LS fits to 2000 Monte Carlo sampled Wiseman plots for nominal ΔH = −6.5 kcal/mol, K = 36,000, $N = 1$. Note that total heat released in each modeled experiment was kept constant, by running all model experiments to 99% saturation of the cell reactant. Left panels: C = 0.5; right panels: C = 50. Top panels: LS fits done with N allowed to float; bottom panels: LS fits done with N fixed at 1. In all cases, the heat error from Eq. (1) was modeled with $\zeta = 0.01$, $\sigma_{irr} = 0.13$ μcal, and 0.6% concentration error in both syringe and cell.

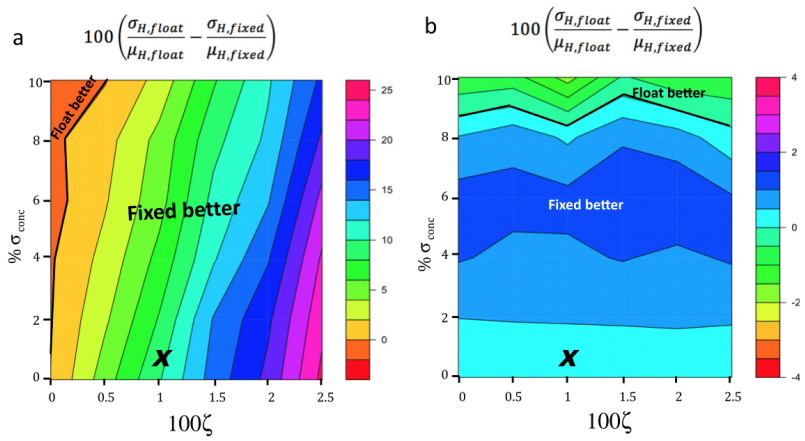


Fig. 7. Contour plots depicting the difference between the percent uncertainties in the fitted enthalpy when N is allowed to float versus being held fixed at 1; i.e., $100(\frac{\sigma_{H,float}}{\mu_{H,float}} - \frac{\sigma_{H,fixed}}{\mu_{H,fixed}})$. Percent heat uncertainty ($\zeta(T)$) is shown along the x-axis, concentration uncertainty is shown along the y-axis. In all cases, the modeled experiments were run to 99% saturation of the cell reactant. a) C value = 0.5; b) C value = 50. Heavy black lines indicate the boundary between the region where fixing N is preferred and that where floating N is preferred. "X" indicates typical experimental conditions, based on the present study. For both experiments, we assumed $\sigma_{irr} = 13\zeta$ µcal.

The contour maps show that it is preferable, under most of the conditions plotted here, to fix N at its nominal value of 1 rather than to let it float. At low C (Fig. 7a), floating N reduces the uncertainty in the enthalpy only under conditions of unrealistically low heat error; while at high C (Fig. 7b), floating N is advantageous only when the concentration errors are greater than ~9%, no matter what the heat error is. Fixing N is clearly advantageous under the typical error conditions found empirically in this study (1% proportional heat error and 0.6% concentration error), which are marked by "x"s in the graphs. For completeness, analogous results for low C experiments run to 55% saturation, rather than 99%, are provided in Supplementary Fig. 6. The region in which fixing N is preferred becomes more dominant in this case.

Because allowing N to float reduces the propagation of cell concentration error while amplifying the propagation of syringe concentration error, it is also of interest to examine the differential effects of these two errors in the context of typical heat noise. Fig. 8 therefore plots $100(\frac{\sigma_{H,float}}{\mu_{H,float}} - \frac{\sigma_{H,fixed}}{\mu_{H,fixed}})$ as a function of syringe and cell concentration errors, with $\zeta = 0.01$ and $\sigma_{irr} = 0.13$ µcal in all cases. At low C (Fig. 8a, note larger scale), fixing N yields lower errors in the enthalpy, except where syringe error is less than ~4% and cell error is greater than ~9%. At optimal C (Fig. 8b), the region where fixing N yields lower error is only slightly larger than that where floating N is preferred; here, to a good approximation, floating N is preferred when the cell error is clearly larger than syringe error, and fixing N is preferred otherwise. Increasing the heat noise to $\zeta = 0.03$ and $\sigma_{irr} = 0.39$ µcal, to model measurements at 65 C°, produces a modest shift of the contours to favor the use of fixed N (Supplementary Fig. 7) in optimal C value plots, while making fixing N favorable everywhere in low C

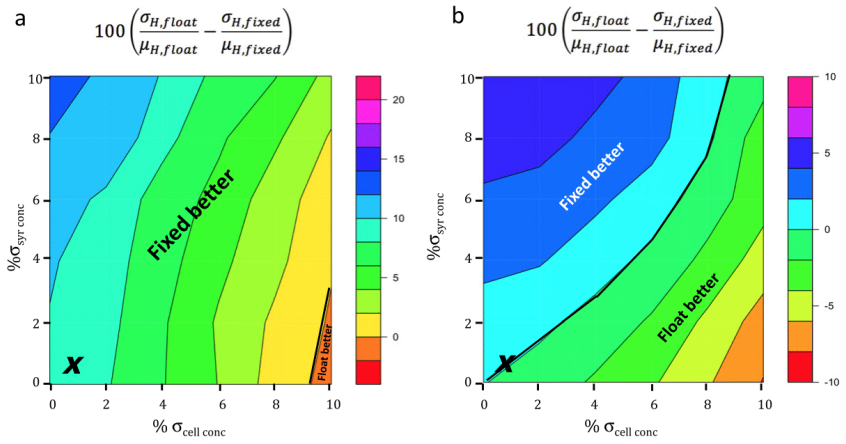


Fig. 8. Same as prior figure except that $\zeta = 0.01$ and $\sigma_{irr} = 0.13$ µcal everywhere and the syringe and cell concentration uncertainties are varied independently.

value plots, consistent with the observation (above) that fixing N reduces the propagation of heat noise to the enthalpy. Again, for completeness, the analogous plot for C = 0.5 and 55% saturation (Supplementary Fig. 8) shows that fixing N is largely advantageous in all examined combinations of syringe and cell error. Again, under normal operating conditions, indicated by the x's, treating N as fixed consistently yields lower errors than allowing it to float.

3.4. Uncertainty Estimates for Realistic Experimental Scenarios

This section applies the MC sampling method, with realistic heat and concentration errors, to compute uncertainties in both binding enthalpy and binding free energy for experimental scenarios spanning optimal and low C values, high and low final saturation of the cell reactant, low and high heat noise, low and high concentration uncertainties, and treatment of N as fixed or floating. In all cases, the experiments modeled use 25 injections. Table 2 provides uncertainties in the enthalpy for modeled experiments in which the uncertainties of the syringe and cell reactants are the same, while Table 3 provides enthalpy results in which these uncertainties are mixed, and Tables 4 and 5 present corresponding uncertainties in the binding free energy. The MC statistics reported are the uncertainty in the fitted value of ΔH or ΔG, the average LS error assigned to ΔH or ΔG across the MC samples, and the standard deviation of N across the MC samples, for cases where it was allowed to float. In all cases, the mean values of ΔH and ΔG deviated minimally from their nominal values (−6.5 kcal/mol and −6.2 kcal/mol, respectively), and the mean value of N, when allowed to float, remained essentially equal to 1, so these quantities are not listed. The key figure of merit in both tables is the error in ΔH (Tables 2,3) and ΔG (Tables 4,5) from MC, as this is the expected

Table 3
Same as prior table, but with mixed combinations of syringe and cell concentration uncertainties.

		Concentration Error			MC Statistics		
	Heat err, ζ	Syr	Cell	N Float/Fixed	ΔH MC Error %	Mean LS ΔH Error %	%SD of N
Optimal C = 50							
	1%	0.6%	1.1%	Fixed	**0.8**	0.4	
	1%	0.6%	1.1%	Float	**0.7**	0.3	1.23
	3%	0.6%	1.1%	Fixed	**1.2**	0.8	
High Sat. (99%)	3%	0.6%	1.1%	Float	**1.3**	0.9	1.36
	1%	1.1%	0.6%	Fixed	**0.6**	0.4	
	1%	1.1%	0.6%	Float	**1.2**	0.3	1.3
	3%	1.1%	0.6%	Fixed	**1.0**	0.8	
	3%	1.1%	0.6%	Float	**1.6**	0.9	1.3
Low C = 0.5							
	1%	0.6%	1.1%	Fixed	**1.3**	0.4	
	1%	0.6%	1.1%	Float	**10.5**	6.2	9.6
	3%	0.6%	1.1%	Fixed	**2.3**	1.2	
High Sat. (99%)	3%	0.6%	1.1%	Float	**36.2**	24.0	26.2
	1%	1.1%	0.6%	Fixed	**0.9**	0.4	
	1%	1.1%	0.6%	Float	**10.6**	6.2	10.0
	3%	1.1%	0.6%	Fixed	**1.9**	1.2	
	3%	1.1%	0.6%	Float	**36.1**	23.6	26.1
	1%	0.6%	1.1%	Fixed	**2.4**	2.1	
	1%	0.6%	1.1%	Float	**26.1**	27.1	18.0
	3%	0.6%	1.1%	Fixed	**6.4**	6.3	
Low Sat. (55%)	3%	0.6%	1.1%	Float	**44.5**	54.2	34.5
	1%	1.1%	0.6%	Fixed	**2.2**	2.1	
	1%	1.1%	0.6%	Float	**26.4**	27.1	17.8
	3%	1.1%	0.6%	Fixed	**6.2**	6.2	
	3%	1.1%	0.6%	Float	**44.4**	54.3	34.3

Table 2
Uncertainties of fitted binding enthalpies for various experimental conditions and assumptions regarding heat and concentration error, where syringe and concentration uncertainties are assumed to be the same. For ζ = 0.01, σ_{irr} = 0.13 μcal; for **ζ** = 0.03, $\boldsymbol{\sigma}_{irr}$ = 0.39 **μ**cal. See text for details.

			Concentration Error			MC Statistics		
		Heat Err, ζ	Syr	Cell	N Float/Fixed	ΔH MC Error %	Mean LS ΔH Error %	%SD of N
Optimal C = 50								
	Low Conc. Error	1%	0.6%	0.6%	Fixed	**0.5**	0.3	
		1%	0.6%	0.6%	Float	**0.7**	0.3	0.8
		3%	0.6%	0.6%	Fixed	**1.0**	0.8	
High Sat. (99%)		3%	0.6%	0.6%	Float	**1.3**	0.9	1
	High Conc. Error	1%	1.1%	1.1%	Fixed	**0.8**	0.4	
		1%	1.1%	1.1%	Float	**1.1**	0.3	1.5
		3%	1.1%	1.1%	Fixed	**1.2**	0.9	
		3%	1.1%	1.1%	Float	**1.6**	0.9	1.6
Low C = 0.5								
	Low Conc. Error	1%	0.6%	0.6%	Fixed	**0.9**	0.4	
		1%	0.6%	0.6%	Float	**10.5**	6.3	9.6
		3%	0.6%	0.6%	Fixed	**1.9**	0.5	
High Sat. (99%)		3%	0.6%	0.6%	Float	**36.2**	23.8	26.1
	High Conc. Error	1%	1.1%	1.1%	Fixed	**1.3**	0.4	
		1%	1.1%	1.1%	Float	**10.5**	6.4	9.6
		3%	1.1%	1.1%	Fixed	**2.1**	1.2	
		3%	1.1%	1.1%	Float	**36.2**	24.0	26.2
	Low Conc. Error	1%	0.6%	0.6%	Fixed	**2.2**	2.1	
		1%	0.6%	0.6%	Float	**26.1**	27.1	18.2
		3%	0.6%	0.6%	Fixed	**6.0**	5.9	
Low Sat. (55%)		3%	0.6%	0.6%	Float	**44.5**	54.4	34.1
	High Conc. Error	1%	1.1%	1.1%	Fixed	**2.5**	2.0	
		1%	1.1%	1.1%	Float	**26.2**	27.3	18.3
		3%	1.1%	1.1%	Fixed	**6.4**	6.3	
		3%	1.1%	1.1%	Float	**44.5**	54.2	34.1

Table 4
Uncertainties of fitted binding free energies for various experimental conditions and assumptions regarding heat and concentration error, where syringe and concentration uncertainties are assumed to be the same. For ζ = 0.01, σ_{irr} = 0.13 **μ**cal; for **ζ** = 0.03, $\boldsymbol{\sigma}_{irr}$ = 0.39 **μ**cal. See text for details.

			Concentration Error			MC Statistics		
		Heat Err, ζ	Syr	Cell	N Float/Fixed	ΔG MC Error %	Mean LS ΔG Error %	%SD of N
Optimal C = 50								
	Low Conc. Error	1%	0.6%	0.6%	Fixed	**0.2**	0.3	
		1%	0.6%	0.6%	Float	**0.2**	0.2	0.8
		3%	0.6%	0.6%	Fixed	**0.6**	0.7	
High Sat. (99%)		3%	0.6%	0.6%	Float	**0.6**	0.6	1.0
	High Conc. Error	1%	1.1%	1.1%	Fixed	**0.4**	0.4	
		1%	1.1%	1.1%	Float	**0.3**	0.3	1.5
		3%	1.1%	1.1%	Fixed	**0.7**	0.8	
		3%	1.1%	1.1%	Float	**0.7**	0.7	1.6
Low C = 0.5								
	Low Conc. Error	1%	0.6%	0.6%	Fixed	**0.2**	0.08	
		1%	0.6%	0.6%	Float	**0.3**	0.1	9.6
		3%	0.6%	0.6%	Fixed	**0.5**	0.1	
High Sat. (99%)		3%	0.6%	0.6%	Float	**1.0**	0.6	26.2
	High Conc. Error	1%	1.1%	1.1%	Fixed	**0.2**	0.08	
		1%	1.1%	1.1%	Float	**0.3**	0.1	9.6
		3%	1.1%	1.1%	Fixed	**0.5**	0.1	
		3%	1.1%	1.1%	Float	**1.0**	0.6	26.2
	Low Conc. Error	1%	0.6%	0.6%	Fixed	**0.3**	0.3	
		1%	0.6%	0.6%	Float	**1.3**	1.1	18.2
		3%	0.6%	0.6%	Fixed	**1.0**	1.0	
Low Sat. (55%)		3%	0.6%	0.6%	Float	**4.3**	4.3	34.1
	High Conc. Error	1%	1.1%	1.1%	Fixed	**0.3**	0.3	
		1%	1.1%	1.1%	Float	**1.4**	1.1	18.3
		3%	1.1%	1.1%	Fixed	**1.0**	1.0	
		3%	1.1%	1.1%	Float	**4.3**	4.3	34.1

Table 5
Same as prior table, but with mixed combinations of syringe and cell concentration uncertainties.

		Concentration Error			MC Statistics		
	Heat err, ζ	Syr	Cell	N Float/Fixed	ΔG MC Error %	Mean LS ΔG Error %	%SD of N
				Optimal C = 50			
	1%	0.6%	1.1%	Fixed	**0.3**	0.4	
	1%	0.6%	1.1%	Float	**0.2**	0.2	1.3
	3%	0.6%	1.1%	Fixed	**0.6**	0.8	
High Sat. (99%)	3%	0.6%	1.1%	Float	**0.6**	0.6	1.4
	1%	1.1%	0.6%	Fixed	**0.3**	0.4	
	1%	1.1%	0.6%	Float	**0.2**	0.3	1.4
	3%	1.1%	0.6%	Fixed	**0.6**	0.6	
	3%	1.1%	0.6%	Float	**0.7**	0.6	1.4
				Low C = 0.5			
	1%	0.6%	1.1%	Fixed	**0.2**	0.08	
	1%	0.6%	1.1%	Float	**0.3**	0.2	9.6
	3%	0.6%	1.1%	Fixed	**0.5**	0.2	
High Sat. (99%)	3%	0.6%	1.1%	Float	**1.0**	0.7	26.2
	1%	1.1%	0.6%	Fixed	**0.2**	0.07	
	1%	1.1%	0.6%	Float	**0.3**	0.2	10.0
	3%	1.1%	0.6%	Fixed	**0.5**	0.2	
	3%	1.1%	0.6%	Float	**1.0**	0.7	28.8
	1%	0.6%	1.1%	Fixed	**0.4**	0.4	
	1%	0.6%	1.1%	Float	**1.3**	1.3	18.0
	3%	0.6%	1.1%	Fixed	**1.2**	1.2	
Low Sat. (55%)	3%	0.6%	1.1%	Float	**4.4**	4.3	34.5
	1%	1.1%	0.6%	Fixed	**0.4**	0.4	
	1%	1.1%	0.6%	Float	**1.4**	1.3	17.8
	3%	1.1%	0.6%	Fixed	**1.2**	1.1	
	3%	1.1%	0.6%	Float	**4.3**	4.2	34.3

uncertainty in the experimental result; the lower this value, the greater the precision of the experiment.

A central result is that experimental error in ΔH is consistently less when N is fixed than when N is allowed to float. The sole exception is the case of high saturation with higher uncertainty in the cell concentration than in the syringe concentration (first two rows of Table 3), and even here floating N improves the uncertainty only from 0.8% to 0.7%. By contrast, there are cases where fixing N improves the uncertainty more than tenfold, relative to floating N. Using Tables 2 and 3, the average ratio of uncertainties from fixing N to uncertainties from floating N is 0.30, indicating the overall improvement of ΔH offered by fixing N. The benefit of fixing N is greatest for the low C experiments. In fact, with N fixed, even a low C experiment with the high 3% heat error associated with measurements at 65 °C (above), can provide results with uncertainties of only ~2%. It is worth noting that the suboptimal 55% saturation low C value plots are more sensitive to error than 99% saturation low C value plots (Tables 2 and 3). However, even here, fixing N can yield reasonable uncertainties, even in 3% heat error scenarios. This observation extends the prior observation [34] that useful precision can be obtained from experiments with low C values. Indeed, it appears that the successes reported there derived less from running the experiments to high saturation, as suggested, and more from the additional step taken in the prior study of fixing N. It is worth mentioning that some experimental uncertainties are below 1%, chiefly in cases where C is optimal. Finally, it is of interest that, when heat and concentration error are combined, as done here, floating N tends to increase the LS error, whereas it was found to reduce the LS error in all cases, when only concentration error was considered (Table 1). More generally, the LS error tends to underestimate uncertainty in ΔH, relative to the more reliable MC error: from Tables 2 and 3, the average ratio of MC error to LS error is 1.7, with minimum of 0.8 and maximum of 3.9.

We also report uncertainties in ΔG for the same combinations of error (Tables 4 and 5). Comparison of Tables 4 and 5 with Tables 2 and 3 shows that the errors in the free energy are smaller than those in the enthalpy, as anticipated. Thus, the ratio of the MC error for ΔH to that for ΔG has a minimum of 1.6, maximum of 36, and mean of 11. We also observe that fixing N is usually advantageous for ΔG, as in the case of ΔH, though not by as much. Thus, for ΔH, the mean ratio of MC error for N fixed versus N float averages 0.30, with minimum and maximum of 0.05 and 1.1, respectively, while for ΔG, the corresponding ratios are 0.7, 0.2, and 1.5. Note that, for both quantities, there are a few instances where fixing N increases uncertainty somewhat, but this weak effect is outweighed by other cases where fixing N is strongly advantageous. Interestingly, the LS error produces more serious underestimates of uncertainty for ΔG than it does for ΔH: the average ratio of MC error to LS error is 1.7, with minimum of 0.6 and maximum of 1.6. This seems to be dependent on the degree of saturation and C value used.

4. Discussion

The present study empirically characterizes two key sources of variance in ITC measurements of binding thermodynamics - the variance in the heat release of each injection, and the variance in the concentration of the syringe solution - and uses Monte Carlo analysis to determine how these uncertainties propagate to the final binding enthalpies under varying different experimental conditions and with different treatments of the stoichiometry during curve-fitting. One general observation is that the manner in which errors in injection heat and concentration propagate to the uncertainty of the binding enthalpy depends on multiple aspects of how the experiments are run, such as the C value and the maximal saturation of the cell reactant. We have tried to explore conditions fairly broadly, and a number of useful and general observations have emerged, as discussed below. Nonetheless, there may be experimentally relevant conditions that were not examined here. So that readers may explore these on their own, and also so they may replicate the present results, we have made the software used here available at (https://github.com/GilsonLabUCSD/Isothermal-Titration-Calorimetry). It is also worth remarking that the present study addresses the precision of the results, not their accuracy; that is, it does not address systematic error. For example, if the calorimeter consistently overestimated heat release by 5%, then all measured enthalpies would be systematically too favorable, but this would not be apparent in the present analysis of variance. We did not probe here for the possibility of such problems with the calibrations of the balance, the volumetric flasks, or the calorimeter. The following subsections summarize the chief results of this study and discuss their significance.

4.1. Advantage of fixing the stoichiometry, N

In many ITC studies, N is allowed to float during the curve fitting process, in order to "absorb" cell concentration error. Indeed, with N floating, cell concentration error does not propagate to ΔH or K, only to N. However, we find that fixing N at its nominal value reduces the uncertainty in ΔH under almost all realistic conditions studied, and this improvement can be dramatic for studies at low C value. Fixing N also tends to reduce the uncertainty in ΔG, but to a lesser extent. It is worth noting that prior studies using global ITC fitting (fitting multiple Wiseman plots) also fix N at an integral number [35,36]. One of the ways that fixing N improves uncertainty in ΔH is by reducing the sensitivity of the fitted Wiseman plot to heat error, $\sigma_Q(T)$. For low C value experiments, where heat error is more problematic, fixing N leads to greatly improved uncertainties. In optimal C value plots, heat error does not propagate to ΔH as severely, but fixing N still leads to small improvements in the uncertainty. The only setting in which allowing N to float may be advantageous is when one knows that the uncertainty in the cell concentration is considerably greater than that of the syringe concentration. For example, if one reactant is a protein that is unstable

or of uncertain purity, it may be advantageous to place this in the cell and allow N to float, or to use a global fitting method in which N is fixed and the concentration uncertainty is explicit. However, in the case of 1:1 binding, where both compounds are robust and well characterized, the uncertainties in their concentrations are normally similar and low. In this setting, fixing $N = 1$ provides better precision than allowing N to float, and this would be our recommended standard procedure.

It has previously been reported that experiments at low C values can yield useful results if they are run to sufficient saturation of the cell reactant [34]. Our results are largely consistent with those. However, we find that the key is not only running to full saturation, but also fixing N. In fact, with $C = 0.5$, even if the experiments are run to a sub-optimal 55% saturation, fixing N affords uncertainties in the binding enthalpy of only a few percent, whereas the uncertainties are much higher when N is allowed to float. It is also worth noting that, although it is often written that error in the syringe concentration propagates proportionally to error in the enthalpy, this is only true when N is allowed to float: fixing N reduces the sensitivity of the enthalpy to the syringe concentration, for experiments at both low and optimal C values. On the other hand, although it has been suggested that fixing N can essentially eliminate the propagation of syringe concentration error to enthalpy error [34], we find that this holds only for experiments run to high saturation at low C values. At low saturation and low C, syringe concentration error does lead to enthalpy error, although less so with N fixed than when N is allowed to float; while, for high C value plots, syringe and cell error both propagate when N is fixed.

Intuitively, it is reasonable that error in the thermodynamic results should be minimized by locking the value of N to its true value (N fixed), rather than allowing it to take on a nonphysical value (N floating). Mathematically, the fact that fixing N typically reduces the uncertainty in ΔH traces to the fact that it reduces the propagation of both syringe concentration error and heat error, and these advantages typically outweigh the increased propagation of cell concentration error. The changes in sensitivity to concentration error upon fixing N trace to the fact that allowing N to float essentially allows the x-axis to be rescaled in a manner that optimizes the fit of the theoretical to the experimental curve, and this can compensate for the effective rescaling along the x-axis that results from any concentration error. Cell concentration error leads to pure x-axis rescaling, and can be completely compensated by allowing N to float, while fixing N removes this compensation mechanism and leaves the fit susceptible to cell concentration error. Errors in the syringe concentration similarly rescale the x-axis, since this is the ratio of the syringe and cell concentrations, but syringe errors also lead to rescaling along the y-axis, as they generate error in the renormalization of the raw heats to molar heats. When N is fixed the effects of these two rescalings on the fitted enthalpy tend to cancel each other, leading to reduced error in the final result. In contrast, when is allowed to float, the x-axes rescaling is corrected by N, leaving the effect of the syringe error on the y-axis uncompensated.

As noted above, the uncertainties in ΔG tend in general to be smaller than the uncertainties in ΔH, but fixing N can still lead to some improvement. The patterns of error appear to be somewhat different from what is observed in ΔH, however, and may warrant additional investigation. For example, there may be cases where fixing N gives the most accurate ΔH, but floating N gives the most accurate ΔG.

4.2. Choosing the number of injections

We find that increasing the number of ITC injections tends to reduce the uncertainty in the binding enthalpy and, to a lesser extent, in the binding free energy. The improvement is marked when C is low and N is allowed to float, and modest when C is normal and/or N is treated as fixed. These results differ from prior results showing lower uncertainties in the case of fewer injections [22]. The difference presumably stems from differences in the experiments modeled and in the assumed magnitude of the heat noise. In particular, the prior study used parameters based on an early-model microcalorimeter [3]; in addition, the author noted that an "artificially large" error was assumed. In contrast, we used the succinic acid/NaOH measurements to develop an empirical error model for the MicroCal VP-ITC, and then used it to model error in typical experiments on the same instrument. From a theoretical standpoint, it has been argued [22,45] that the relative error (noise-to-signal ratio) of each injection will rise when more injections are used, as each injection becomes smaller. However, this holds only when the fixed component of the heat noise is dominant; for the experiments modeled here, most of the heat error is proportional in character (Eq. (1)), so the relative error of each injection does not rise much as the number of injections is increased. Under these conditions, there is no significant penalty for dividing the experiment across more injections, and doing so may improve precision by providing a more detailed delineation of the binding isotherm. Of course, if the heat is divided so much that the proportional part of the error falls below the irreducible part of the error, then every peak will have a similar fixed error, likely leading to the same conclusions observed previously.

Thus, it would appear optimal to maximize the number of injections, subject to the constraint that the smaller injections still yield heats for which the proportional error is greater than or equal to the irreducible error. That said, experiments with good C values and/or low C values and fixed N can yield excellent results with only a few (e.g., 10–15) injections. In such cases, the modest increase in uncertainty from reducing the number of injections may well be acceptable. Finally, it should be noted that, for a different calorimeter, the fixed heat error might dominate over the proportional error, in which case, fewer injections might indeed be preferable.

4.3. Variance in injection heat

Using the succinic acid experiment described in the Results, we found that the heat error may be described as a combination of a term proportional to the amount of heat released per injection, and a small irreducible error (Eq. (1)), and that there is low correlation across injections during an experiment. The constant of proportionality, $\zeta(T)$, is about 0.01 at room temperature and rises to 0.03 at 65 °C; and the irreducible error σ_{irr}, similarly rises threefold over this temperature range, from 0.13 to 0.39 μcal/injection. We are not aware of any prior report that the heat uncertainty depends upon the temperature. It should be emphasized that these results are specific to the VP-ITC instrument we studied, and different results could be obtained for different calorimeters. Indeed, a prior study of this type for the iTC200 instrument does not show a clear pattern of proportional error [42], although the scatter in the data points makes this difficult to pin down. Fortunately, the succinic acid experiment is straightforward and the materials are inexpensive, so it may easily be used to characterize the heat error for whatever instrument one is using.

Our observation that the uncertainty in the injection heat is mainly proportional to the heat evolved, and is not dependent on injection volume, runs counter to models that the heat error is independent of the injection heat [33,37]; or that, that, for the VP-ITC, the heat error comprises a large, constant heat-measurement error plus a constant volume-delivery error [22]. Although the latter model is based on uncertainties reported by Wiseman and coworkers [3], Wiseman and coworkers did not describe how the heat uncertainty depended on the heat of injection or the volume, so it is not clear whether their uncertainties conformed to this particular model. In addition, their study was apparently carried out on a pioneering early calorimeter, rather than the commercial VP-ITC unit, and it is probable that the two instruments have quite different patterns of error.

Importantly, we observe much larger values of the heat noise, for typical injections, than the value of 0.1 μcal previously suggested as a

simple constant model. For example, a 1% error on a 200 μcal heat of injection will result in a roughly 2 μcal error. It is also worth mentioning that, although a prior error model for the heat error also include both a proportional part and an irreducible part [34], instead of adding the variances in order to model two independent noise processes, they set the irreducible standard deviation as a floor to a proportional component with a temperature-independent coefficient:

$$\sigma_Q \approx \max(\zeta Q, \sigma_{irr}) \quad (2)$$

Estimating overall error by adding the variances [38], as done here (Eq. (1), would seem to be on a more solid statistical footing and also gives a somewhat more conservative (i.e., larger) estimate of the uncertainty. However, the small magnitude of the irreducible error makes this distinction relatively unimportant. Finally, the data in Fig. 2c would argue against a prior model in which the non-fixed component of the heat error was assumed to result from variance in the injection volumes [11], as reducing concentration while keeping injection volumes constant shows the same proportional trend.

A minor note of interest is the possible correlation between successive injection heats seen in our results. Although the R^2 is small and variable across datasets, there still seems to be some positive correlation. This could arise from baseline drift, where the baseline steadily moves in a particular direction, so that adjacent peaks have similar biases. This might be correctable by using software that more rigorously analyzes baselines [24,46]. Regardless, the effect seems small and unlikely to significantly influence the uncertainty in the thermodynamic results.

4.4. Concentration errors

It has been suggested that the high variance in binding enthalpies in the ABRF-MIRG'02 study result largely from variance in the concentration of the syringe reactant, 4-carboxybenzenesulfonamide (CBS) [32]. This attribution is based on the fact that the participants' data measured apparent extinction coefficients of CBS ranging over at least 1100–1400 L/mol/cm, a range of about 25%. The participants had been provided with powdered CBS, and thus made up their own solutions. Nonetheless, we would not have expected that this procedure would yield such large errors, so we ran a simple empirical test, using β-CD as the solute, and found concentration uncertainties on the order of 1%, which is more consistent with the nominal precision of a typical balance and flask. Given that CBS was shown to be stable [32], it is not clear how 20-fold larger concentration errors could have arisen in the prior study, even if some participants were new to ITC, and we would be cautious about assuming that such high concentration errors are typical, particularly as there was no direct verification of their existence.

4.5. Experimental uncertainties in binding thermodynamics measured by ITC

It has been noted previously that the LS method of estimating uncertainties does not account for concentration error [26]. However, we find that, with N fixed, the LS uncertainty can to some degree capture the consequences of concentration error. On the other hand, we also find that LS uncertainty tends in general to underestimate the actual uncertainties, due to both types of error, as measured by comparison with reference MC calculations. Thus, the LS uncertainties provided by commonly used software packages can significantly underestimate the true uncertainty, though the degree of this underestimate depends on experimental factors, such as the C value and whether N is allowed to float or treated as fixed. We also observed that low C value Wiseman plots run to a sub-optimal 55% saturation are much more sensitive t o heat noise and concentration error than low C value plots run to 99% saturation, as evidenced in Tables 2 and 3. However, fixing N allows even the 3% heat error cases with low C values and 55% saturation to provide enthalpies with uncertainties of only ~2–6 %, and free energies with uncertainties of only ~0.3–1%.

Unexpectedly, low C value Wiseman plots run to a sub-optimal 55% saturation are less sensitive to heat error than plots run to 99% saturation, as evidenced in Tables 2 and 3. This may be the result of the lack of data points at the beginning of the curve; as seen in Supplementary Fig. 2, 99% saturation, low C value plot rapidly rises to saturation, with only three or four points along the steep rise of the curve. This phenomenon deserves further study, and might be corrected for by using more injections of smaller volume, providing more points in the rising part of the curve. However, low C value Wiseman plots with sub-optimal saturation show slightly higher uncertainties in fitted ΔG, as evidenced in Tables 4 and 5.

Under some realistic experimental conditions and error assumptions, we observe small systematic errors in the reported values of the binding enthalpy, for reasons that are unclear. However, this bias arises only when N is allowed to float; and, when observed, the bias only exceeds 5% at around 3% heat error, and, in this regime, the uncertainty in the binding enthalpy is several times larger than the bias and for most purposes effectively drowns it out.

The present study suggests that, with a MicroCal VP in good working condition, for optimal C value experiments at room temperature, where solutions are made up in larger volumetric flasks, it is reasonable to expect uncertainties of 0.5% in the binding enthalpy and 0.2% in the binding free energy. If one is forced to work with smaller volumetrics, due, for example, to limited compound availability, the uncertainties in enthalpy and free energy may rise to on the order of 1–2% and 0.4–0.7%, respectively. At lower C values, which may be difficult to avoid in the case of weak binding and/or low solubility, one may still anticipate keeping uncertainties in enthalpy below about 2%, rising to about 4% for measurements at 65 °C. The corresponding values for free energies are 0.2% and 0.5% Note that these values depend on keeping N fixed; floating N allows the uncertainties for the same experiments to approach 50% in the most challenging cases considered here. Thus, so long as N is held fixed during curve-fitting, it should be possible to obtain quite precise thermodynamic data on a wide range of binding systems at various temperatures.

5. Conclusions

In summary, the present study yields several broad conclusions regarding the design and interpretation of isothermal titration microcalorimetry experiments.

- For the instrument studied here, the standard deviation of the heat of injection is approximately proportional to the integrated heat of the injection, with proportionality coefficients of 1% at room temperature and 3% at 65 °C. The proportional error is accompanied by a small irreducible uncertainty of about 0.13 μcal at room temperature and 0.39 μcal at 65 °C.
- Experimental evaluation of concentrations of solutions made in our laboratory indicate a 0.6–1.1% uncertainty, depending on the volume of solution made up.
- Fixing N alleviates the propagation of heat error to the binding enthalpy, especially for low C value Wiseman plots.
- Concentration errors propagate depending on how N is treated: floating N allows a near-linear propagation of syringe error to fitted enthalpies of binding; fixing N introduces sensitivity to cell error while simultaneously decreasing sensitivity to syringe error. When one fixes N in fitting low C value experiments, the sensitivity of the enthalpy to the syringe concentration error depends strongly on the degree of saturation to which the experiment is run.
- Under a wide range of experimental conditions and designs, fixing N at its nominal value yields less uncertainty in the binding enthalpy and free energy than does allowing N to float. This is particularly true for low C value experiments. Of the conditions examined here,

floating N is preferable only when the concentration uncertainty of the cell reactant is high, and particularly when the experiment is done at a high C value.
- Experimental uncertainties are typically considerably lower for binding free energies than for binding enthalpies.
- Error estimates from least squares fitting tend to underestimate the more realistic error estimates obtained by Monte Carlo analysis, particularly for enthalpies measured based on well-saturated Wiseman plots, at either optimal or low C values.
- Given stable and well-characterized reagents, and with N treated as fixed, it is realistic to expect uncertainties in binding enthalpies within about 2% for ITC measurements under a wide range of experimental conditions, with lower uncertainties for the corresponding binding free energies
- The present study does not address all conditions that may arise in practice, and the software used in the present study may be utilized to explore other experimental setups of interest.

Transparency Document

The Transparency document associated with this article can be found in online version.

Acknowledgements and Disclosures

We are grateful to Drs. Elizabeth Komives, Deepa Balasubramaniam, and Verna Frasca for generously sharing their expertise and answering many questions. M.K.G. has an equity interest in, and is a cofounder and scientific advisor of VeraChem LLC. S.A.K. acknowledges training support from the Molecular Biophysics training grant (T32 GM008326). This publication was supported by the National Institute of General Medical Sciences of the National Institutes of Health (NIH) Grant GM61300. These findings are solely of the authors and do not necessarily represent the views of the NIH.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.bbagen.2016.09.002.

References

[1] P. Englebienne, A.V. Hoonacker, M. Verhas, Surface plasmon resonance: principles, methods and applications in biomedical sciences, J. Spectrosc. 17 (2003) 255–273.
[2] A. Zettner, Principles of Competitive Binding Assays (Saturation Analyses). I. Equilibrium Techniques, Clin. Chem. 19 (1973) 699–705.
[3] T. Wiseman, S. Williston, J.F. Brandts, L.-N. Lin, Rapid measurement of binding constants and heats of binding using a new titration calorimeter, Anal. Biochem. 179 (1989) 131–137.
[4] J.E. Ladbury, B.Z. Chowdhry, Sensing the heat: the application of isothermal titration calorimetry to thermodynamic studies of biomolecular interactions, Chem. Biol. 3 (1996) 791–801.
[5] S. Leavitt, E. Freire, Direct measurement of protein binding energetics by isothermal titration calorimetry, Curr. Opin. Struct. Biol. 11 (2001) 560–566.
[6] A. Velazquez-Campoy, S.A. Leavitt, E. Freire, Characterization of protein–protein interactions by isothermal titration calorimetry, Methods Mol. Biol. Clifton N. J. 261 (2004) 35–54.
[7] M.W. Freyer, E.A. Lewis, Methods in Cell Biology, 84, Elsevier, 2008 79–113.
[8] G.A. Holdgate, Making cool drugs hot: isothermal titration calorimetry as a tool to study binding energetics, BioTechniques 31 (2001) 164–166 (168, 170 passim).
[9] P. Gilli, V. Ferretti, G. Gilli, P.A. Borea, Enthalpy-entropy compensation in drug-receptor binding, J. Phys. Chem. 98 (1994) 1515–1518.
[10] J.E. Leffler, The Enthalpy-Entropy Relationship and its Implications for Organic Chemistry, J. Org. Chem. 20 (1955) 1202–1231.
[11] R. Lumry, S. Rajender, Enthalpy-entropy compensation phenomena in water solutions of proteins and small molecules: a ubiquitous property of water, Biopolymers 9 (1970) 1125–1227.
[12] T.S.G. Olsson, J.E. Ladbury, W.R. Pitt, M.A. Williams, Extent of enthalpy-entropy compensation in protein–ligand interactions, Protein Sci. Publ. Protein Soc. 20 (2011) 1607–1618.
[13] C. Krintel, et al., Enthalpy-Entropy Compensation in the Binding of Modulators at Ionotropic Glutamate Receptor GluA2, Biophys. J. 110 (2016) 2397–2406.
[14] G.M. Whitesides, V.M. Krishnamurthy, Designing ligands to bind proteins, Q. Rev. Biophys. 38 (2005) 385–395.
[15] P.C. Weber, F.R. Salemme, Applications of calorimetric methods to drug discovery and the study of protein interactions, Curr. Opin. Struct. Biol. 13 (2003) 115–121.
[16] J.-P. Ryckaert, A. Bellemans, Molecular dynamics of liquid alkanes, Faraday Discuss. Chem. Soc. 66 (1978) 95–106.
[17] W.L. Jorgensen, Quantum and statistical mechanical studies of liquids. 10. Transferable intermolecular potential functions for water, alcohols, and ethers. Application to liquid water, J. Am. Chem. Soc. 103 (1981) 335–340.
[18] W.L. Jorgensen, D.S. Maxwell, J. Tirado-Rives, Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids, J. Am. Chem. Soc. 118 (1996) 11225–11236.
[19] S. Lifson, A.T. Hagler, P. Dauber, Consistent force field studies of intermolecular forces in hydrogen-bonded crystals. 1. Carboxylic acids, amides, and the C: O.cntdot..cntdot.cntdot.H- hydrogen bonds, J. Am. Chem. Soc. 101 (1979) 5111–5121.
[20] A.T. Fenley, N.M. Henriksen, H.S. Muddana, M.K. Gilson, Bridging Calorimetry and Simulation through Precise Calculations of Cucurbituril–Guest Binding Enthalpies, J. Chem. Theory Comput. 10 (2014) 4069–4078.
[21] N.M. Henriksen, A.T. Fenley, M.K. Gilson, Computational Calorimetry: High-Precision Calculation of Host–Guest Binding Thermodynamics, J. Chem. Theory Comput. 11 (2015) 4377–4394.
[22] J. Tellinghuisen, A study of statistical error in isothermal titration calorimetry, Anal. Biochem. b321 (2003) 79–88.
[23] J. Tellinghuisen, Optimizing experimental parameters in isothermal titration calorimetry, J. Phys. Chem. B 109 (2005) 20027–20035.
[24] S. Keller, et al., High-Precision Isothermal Titration Calorimetry with Automated Peak Shape Analysis, Anal. Chem. 84 (2012) 5066–5073.
[25] J. Tellinghuisen, J.D. Chodera, Systematic errors in isothermal titration calorimetry: concentrations and baselines, Anal. Biochem. 414 (2011) 297–299.
[26] S.E. Boyce, J. Tellinghuisen, J.D. Chodera, Avoiding accuracy-limiting pitfalls in the study of protein-ligand interactions with isothermal titration calorimetry, bioRxiv (2015) 23796, http://dx.doi.org/10.1101/023796.
[27] B.A. Pethica, Misuse of thermodynamics in the interpretation of isothermal titration calorimetry data for ligand binding to proteins, Anal. Biochem. 472 (2015) 21–29.
[28] L. Baranauskiene, V. Petrikaite, J. Matuliene, D. Matulis, Titration calorimetry standards and the precision of isothermal titration calorimetry data, Int. J. Mol. Sci. 10 (2009) 2752–2762.
[29] M.M. Pierce, C.S. Raman, B.T. Nall, Isothermal titration calorimetry of protein–protein interactions, Methods 19 (1999) 213–221.
[30] A.A. Travers, M. Buckle, DNA-protein Interactions: A Practical Approach, Oxford University Press, 2000.
[31] R.N. Goldberg, N. Kishore, R.M. Lennen, Thermodynamic Quantities for the Ionization Reactions of Buffers, J. Phys. Chem. Ref. Data 31 (2002) 231–370.
[32] D.G. Myszka, et al., The ABRF-MIRG'02 study: assembly state, thermodynamic, and kinetic analysis of an enzyme/inhibitor interaction, J. Biomol. Tech. 14 (2003) 247–269.
[33] M.E. Harrous, A. Parody-Morreale, Measurement of biochemical affinities with a Gill titration calorimeter, Anal. Biochem. 254 (1997) 96–108.
[34] W.B. Turnbull, A.H. Daranas, On the value of c: can low affinity systems be studied by isothermal titration calorimetry? J. Am. Chem. Soc. 125 (2003) 14859–14866.
[35] J.C.D. Houtman, et al., Studying multisite binary and ternary protein interactions by global analysis of isothermal titration calorimetry data in SEDPHAT: application to adaptor protein complexes in cell signaling, Protein Sci. Publ. Protein Soc. 16 (2007) 30–42.
[36] H. Zhao, G. Piszczek, P. Schuck, SEDPHAT–a platform for global ITC analysis and global multi-method analysis of molecular interactions, Methods 76 (2015) 137–148.
[37] V.H. Le, R. Buscaglia, J.B. Chaires, E.A. Lewis, Modeling Complex Equilibria in ITC Experiments: Thermodynamic Parameters Estimation for a Three Binding Site Model, Anal. Biochem. 434 (2013) 233–241.
[38] J. Tellinghuisen, Statistical error in isothermal titration calorimetry: Variance function estimation from generalized least squares, Anal. Biochem. 343 (2005) 106–115.
[39] J. Tellinghuisen, Optimizing isothermal titration calorimetry protocols for the study of 1:1 binding: Keeping it simple, Biochim. Biophys. Acta (2015), http://dx.doi.org/10.1016/j.bbagen.2015.10.011.
[40] J.D. Chodera, D.L. Mobley, Entropy-enthalpy compensation: role and ramifications in biomolecular ligand recognition and design, Annu. Rev. Biophys. 42 (2013) 121–142.
[41] Z. Bikádi, R. Kurdi, S. Balogh, J. Szemán, E. Hazai, Aggregation of cyclodextrins as an important factor to determine their complexation behavior, Chem. Biodivers. 3 (2006) 1266–1278.
[42] W.B. Peters, V. Frasca, R.K. Brown, Recent developments in isothermal titration calorimetry label free screening, Comb. Chem. High Throughput Screen. 12 (2009) 772–790.
[43] J. Carrazana, A. Jover, F. Meijide, V.H. Soto, J. Vazquez Tato, Complexation of adamantyl compounds by beta-cyclodextrin and monoaminoderivatives, J. Phys. Chem. B 109 (2005) 9719–9726.
[44] VP-ITC MicroCalorimeter User's Manual, Rev. E.
[45] J. Tellinghuisen, Stupid Statistics!. in Biophysical Tools for Biologists: In Vitro Techniques, in: J.J. Correia, H.W. Detrich III (Eds.), Methods in Cell Biology, v. 84, Elsevier, London, 2008.
[46] T.H. Scheuermann, C.A. Brautigam, High-precision, automated integration of multiple isothermal titration calorimetric thermograms: new features of NITPIC, Methods 76 (2015) 87–98.

Chapter 4, in full, is a reprint of the material as it appears in “Evaluation and minimization of uncertainty in ITC binding measurements: heat error, concentration error, saturation, and stoichiometry”, in SM Kantonen, NM Henriksen, MK Gilson. Biochimica et Biophysica Acta (BBA)-General Subjects (2017). The dissertation author was primary investigator and author of this publication.

5 ACCOUNTING FOR APPARENT DEVIATIONS BETWEEN CALORIMETRIC AND VAN'T HOFF ENTHALPIES

BBA - General Subjects 1862 (2018) 692–704

Contents lists available at ScienceDirect

BBA - General Subjects

journal homepage: www.elsevier.com/locate/bbagen

Accounting for apparent deviations between calorimetric and van't Hoff enthalpies

Samuel A. Kantonen, Niel M. Henriksen, Michael K. Gilson*

Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0736, USA

ARTICLE INFO

Keywords:
ITC
van't Hoff
Binding thermodynamics
Uncertainty

ABSTRACT

Background: In theory, binding enthalpies directly obtained from calorimetry (such as ITC) and the temperature dependence of the binding free energy (van't Hoff method) should agree. However, previous studies have often found them to be discrepant.
Methods: Experimental binding enthalpies (both calorimetric and van't Hoff) are obtained for two host-guest pairs using ITC, and the discrepancy between the two enthalpies is examined. Modeling of artificial ITC data is also used to examine how different sources of error propagate to both types of binding enthalpies.
Results: For the host-guest pairs examined here, good agreement, to within about 0.4 kcal/mol, is obtained between the two enthalpies. Additionally, using artificial data, we find that different sources of error propagate to either enthalpy uniquely, with concentration error and heat error propagating primarily to calorimetric and van't Hoff enthalpies, respectively.
Conclusions: With modern calorimeters, good agreement between van't Hoff and calorimetric enthalpies should be achievable, barring issues due to non-ideality or unanticipated measurement pathologies. Indeed, disagreement between the two can serve as a flag for error-prone datasets. A review of the underlying theory supports the expectation that these two quantities should be in agreement.
General significance: We address and arguably resolve long-standing questions regarding the relationship between calorimetric and van't Hoff enthalpies. In addition, we show that comparison of these two quantities can be used as an internal consistency check of a calorimetry study.

1. Introduction

The thermodynamics of molecular recognition is of central importance in a range of biological and biomedical fields, spanning molecular biophysics to drug design. The thermodynamic parameter of primary interest is the standard binding free energy, ΔG^o, as it determines the binding affinity; but the standard binding enthalpy, ΔH^o, and entropy, ΔS^o, are also of interest, as they further characterize the molecular processes associated with binding, and can be used as reference data to test and improve the accuracy of computational molecular simulations [1]. Until about 25 years ago, the chief approach to measuring the binding enthalpy was to measure the binding free energy at several temperatures and analyze the resulting data with the van't Hoff relation [2],

$$\Delta H^o = RT^2\left(\frac{d \ln K}{dT}\right)_p \quad (1)$$

where K is the binding constant, R is the gas constant, T is absolute temperature, and the derivative is taken at constant pressure. The situation changed with the introduction of isothermal titration calorimetry (ITC) instrumentation [3], which directly determines not only ΔG^o but also ΔH^o from measurements at a single temperature.

Somewhat unexpectedly, however, ΔH^o values obtained from the van't Hoff method and from the direct calorimetric method for the same binding reaction were early reported to be inconsistent in several cases [4–6], with deviations of up to 10 kcal/mol. Building on these discrepancies, it was argued that the theory underlying the van't Hoff expression might not apply in all cases, so that inconsistencies between the direct and van't Hoff enthalpies should be expected [7,8]. However, this view was vigorously disputed [9,10], and at least one experimental study showed that apparently large inconsistencies (up to ~3 kcal/mol) actually could be statistically insignificant [11]. Indeed, one limitation of this experimental study was the use of a relatively early model calorimeter, which lacked the precision of more modern instruments (J.R. Horn, personal communication), and thus could not provide a strong

Abbreviations: ITC, isothermal titration calorimetry; LS, least squares; MC, Monte Carlo; VH, van't Hoff

* Corresponding author.
E-mail address: mgilson@ucsd.edu (M.K. Gilson).

https://doi.org/10.1016/j.bbagen.2017.11.020
Received 14 July 2017; Received in revised form 15 November 2017; Accepted 28 November 2017
Available online 06 December 2017

test of the consistency between the values of enthalpy from the van't Hoff and direct calorimetric methods. More recent ITC examinations of this issue have used higher-precision instruments. In one, statistically significant deviations between van't Hoff and direct calorimetric enthalpy values were tentatively ascribed to procedural issues, rather than to inapplicability of the van't Hoff relation [6]. In another, though, a new apparent experimental violation of the van't Hoff relation was demonstrated experimentally, and it was again argued on theoretical grounds that consistency between the two values of the binding enthalpy should not be expected [12]. However, both of these more recent studies involved injection of dications, Ba^{2+} or Ca^{2+}, into the reaction cell, an experimental design that could conceivably have generated temperature-dependent changes in activity coefficients and partial molar enthalpies that might complicate interpretation of the data.

Here, we reexamine this issue with new experimental tests of the consistency between van't Hoff and directly measured enthalpy values, for ITC experiments that reduce the potential for changes in solution ideality during the experiment by using an electrically neutral molecular host, β-cyclodextrin, as the titrant. Such compact molecular recognition systems are finding increasing application as models to evaluate and improve the reliability of computational models of binding [13–15], and we are particularly interested in assessing the reliability of binding enthalpies, because computational methods of obtaining these from simulations now offer the possibility of comparing not only binding free energies, but also enthalpies, with experiment [14,15]. In addition, a number of prior studies that have reported discrepancies between van't Hoff and direct calorimetric enthalpies also focused on host-guest systems, such as crown ethers and cyclodextrins [4–6,12]. However, the basic results are expected to be applicable not only to other host-guest systems, but also to biomolecules. The experiments are carried out on a modern instrument with a low level of uncertainty in its heat measurements [16]. We furthermore use mathematical modeling to investigate how two common sources of experimental error, concentration error and heat error [16], can generate apparent violations of the van't Hoff relation. We find that, although van't Hoff enthalpies are sensitive to errors in the binding free energy, they may nonetheless be expected to agree with directly measured enthalpies to within about 10%, given typical experimental errors. We also find that concentration error chiefly affects the direct enthalpy, while heat error chiefly affects van't Hoff enthalpies, and that marked inconsistency between the van't Hoff and direct enthalpies can be used to flag pathological measurements. We close by discussing the implications of these results and considering recent theoretical arguments against the applicability of the van't Hoff relation to aqueous binding thermodynamics.

2. Materials and methods

2.1. Materials

β-cyclodextrin (catalog no. C-4767), rimantadine hydrochloride (catalog no. 390593), and amantadine hydrochloride (catalog no. 138576) were obtained from Sigma-Aldrich Company (St. Louis, MO). Previous older batches (over one year after purchase) of β-cyclodextrin showed some evidence of aggregation in newly made solutions that had stood for on the order of an hour or more, based on slight clouding of concentration solutions and spectrophotometric detection, so a new lot was used for experiments presented in this study. Solutions made from the new lot showed no visible evidence of aggregation. NMR spectra were taken for β-cyclodextrin, rimantadine, and amantadine to verify structure and purity.

2.2. Isothermal titration calorimetry experiments

ITC experiments were performed using a MicroCal model VP-ITC (MicroCal, Northampton, CT, Serial Number 01-08-930). Previous work has emphasized that even minor deflections in the power baseline during ITC experimentation can lead to nontrivial errors that, when feasible, requires detailed analysis to correct. We therefore investigated sources of baseline noise for our experimental setup. For our VP-ITC instrument, we found that small, deliberate vibrations of the laboratory bench on which the calorimeter was set caused significant deflections in baseline. To prevent this, the calorimeter was placed in an isolated room on a 2″ thick block of urethane foam. This setup eliminated deflections during deliberate vibration of the bench and provided a more stable baseline. Additionally, a purpose-built, clear acrylic shield was used to reduce possible temperature shifts due to drafts. The instrument's built-in Y-axis calibrations were performed, at 27 °C, to verify that the instrument was responding to known power inputs within the 1% error tolerance prescribed by the manufacturer.

Solutions of β-cyclodextrin, rimantadine, and amantadine were prepared in 10 mM phosphate buffered saline (pH 7.4). In all experiments, β-cyclodextrin was titrated into the cell, in order to take advantage of both its negligible heat of dilution (Supplementary Fig. 1) and its electrical neutrality, which reduces the likelihood of large changes in activity coefficients and partial molar enthalpies in the course of the experiment. Because the measured binding enthalpy is particularly sensitive to the concentration of the syringe titrant, the β-cyclodextrin solutions were prepared in relatively large quantities, as this reduces errors in the concentration due to weighing errors. The solutions thus were prepared in 250 mL volumetric flasks at concentrations of 12.8–13.4 mM, which corresponds to about 3 g of β-cyclodextrin per 250 mL batch of solution. The concentration of both guests was 1.1–1.4 mM. This is the maximum concentration of β-cyclodextrin attainable in aqueous buffer. Solutions of the guest molecules in the ITC cell were prepared in 25 mL volumetric flasks. Masses of both host and guest were measured using a Sartorius CPA225D Micro Balance. An additional set of experiments for each guest was performed with reduced concentrations (0.27 mM rimantadine and 0.23 mM amantadine, with 5.1–6 mM β-cyclodextrin), using the same procedures described above, to check for evidence of nonideal solution properties. For each guest, duplicate experiments were performed, and fresh solutions were prepared for each replicate. However, a single set of cell and syringe solutions was used for each complete temperature series, to keep the concentrations precisely the same across the measurements in the series.

ITC experiments were designed to allow maximum signal while staying within the instrument's limitations in measured power, and thus to maximize the signal-to-noise ratio. Thus, at each temperature, the reference power was set as low as possible without allowing overshoot due to generation of binding heat at a rate in excess of the reference power. For amantadine, 27 10-uL injections were performed at each temperature. For rimantadine, runs were done using 54 5-uL injections and 27 10-uL injections. Smaller injection sizes were chosen for rimantadine because of its more enthalpic reactions; for the concentrations chosen for rimantadine, the raw heat signal from 10 μL injections overwhelmed the calorimeter, causing problems with the baseline. For all experiments, the first injection was discarded; this practice is ubiquitous in ITC experimentation, as the heat of the first injection is often inappropriately small, due to diffusive mixing of cell and syringe solutions within the tip of the syringe prior to the first injection. The time between injections for a typical experiment in this study was extended from the default of 120 s to 300 or 500 s, to allow for a more complete return to baseline.

2.3. Generation of artificial ITC data

We generated and analyzed artificial Wiseman plots based on thermodynamic parameters for the reversible association of a model system previously studied by both our laboratory and others: the adamantyl-based drug rimantadine with the cyclic oligosaccharide β-cyclodextrin. The data were modeled for our own experimental setups

with rimantadine in the cell and cyclodextrin in the syringe, with binding thermodynamics and Wiseman plot C values similar to experimental conditions from 300 K to 340 K (50 to 6, respectively). All modeled data use 25 injections of 10 μL to generate Wiseman plots. Following our prior investigation into heat error for this instrument [16], we estimated the standard deviation of the measured heat of a given injection as

$$\sigma_Q \approx \{[\zeta(\mathrm{T})Q]^2 + \sigma_{irr}^2(\mathrm{T})\}^{\frac{1}{2}} \quad (2)$$

where σ_Q is the standard deviation of the measured injection heat; ζ(T) is the temperature dependent coefficient of the proportional error component; and σ_{irr}(T) is the standard deviation of the temperature-dependent irreducible heat noise. Based on our prior study [16], we set ζ to 0.01 (1%) for T = 27–57 °C, and to 0.03 (3%) for experiments at 67 °C. Similarly, σ_{irr} was set to 0.13 μcal and 0.39 μcal for experiments at 27–57 °C and at 67 °C, respectively.

We also used model ITC data to determine whether thermal expansion of the syringe and cell solutions, as T is increased from 300 K to 340 K to generate data for the van't Hoff method, might lower concentrations enough to materially affect ΔH_{Direct} and ΔH_{VH}. Accounting for the ~1.4% thermal expansion over this temperature had very little effect on the results (Supplementary Table 2), so we did not include thermal expansion in subsequent modeling or data analysis.

2.4. Analysis of ITC data

The Origin 7.0 software was used to process the raw experimental data by integrating it to give the heat release per injection. Using custom Python scripts, either artificial model data or the experimental raw heats were normalized by concentration of injectant to produce a Wiseman plot, and values of K and ΔH for the binding reactions were determined by using non-linear optimization (Marquardt-Levenberg) with theoretical equations described by Wiseman [3]. The values of the binding enthalpy from this procedure are termed the direct results, ΔH_{Direct}, while values of the binding enthalpy obtained from the van't Hoff relationship (see below) are termed van't Hoff results, ΔH_{VH}. (Although the superscript "o" is omitted, for simplicity, both quantities pertain to the usual standard condition where the reactants and product are present in a hypothetical ideal 1 M solution.)

We used bootstrap analysis to study how heat error and concentration error propagate to the fitted quantities ΔH_{Direct} and ΔH_{VH}, for both the experimental and simulated data. For each point on the Wiseman plot, heat error was modeled by sampling the heat of each injection from a Gaussian with a mean of the measured or simulated heat, and standard deviation σ_Q, from the equation above. These Gaussian samples allowed the construction of multiple artificial Wiseman plots, each with its data points sampled from its respective Gaussian. Additionally, each artificial Wiseman plot was recalculated multiple times with various concentration errors assumed. Importantly, the same concentration was always used for each full set of Wiseman plots across temperatures, to simulate doing a set of temperature points measured for the same solution, as done in our experiments. Thus, although there is some error in the concentration of the solution, the same solution is used at all temperatures to get an internally consistent set of binding free energies for use in the van't Hoff analysis.

Values of K and ΔH_{Direct} were fitted to each resulting Wiseman plot, to yield a statistical distribution of these fitted quantities, allowing calculation of the mean and standard deviation of K and ΔH_{Direct} across the bootstrapped samples. This bootstrapping can more fully capture the uncertainties in the reported thermodynamic quantities than the common approach, used in Origin and other programs, of reporting errors based only on the fit to a single Wiseman plot [16]. The overall mean and SEM of the ΔH_{VH} and K values for each temperature point, across experimental duplicates, were determined by taking the mean and standard deviation over all replicates, from bootstrapping and from the duplicated measurements. Thus,

$$\overline{\mathrm{K}} = \frac{\sum_{i=1}^{n} K_{i,1} + \sum_{i=1}^{n} K_{i,2}}{2n} \quad (3)$$

$$\sigma_{\mathrm{K}} = \sqrt{\frac{\sum_{i=1}^{n} (K_{i,1} - \overline{\mathrm{K}})^2 + \sum_{i=1}^{n} (K_{i,2} - \overline{\mathrm{K}})^2}{2n}} \quad (4)$$

where $K_{i,1}$ is the fitted value from the i^{th} Wiseman plot during bootstrapping, over n Wiseman plots based on the first measurement, and $K_{i,2}$ is the fitted value from the i^{th} Wiseman plot during bootstrapping, over n Wiseman plots based on the duplicate measurement. The mean and standard deviation of ΔH_{VH} was determined with analogous formulae. This bootstrapping process realistically models the propagation of uncertainties in the raw data to the derived thermodynamic properties. For model Wiseman plots, the data were processed both with the stoichiometry, N, fixed at 1, since one-to-one binding was used to generate the simulated data; and with N allowed to float as an additional adjustable parameter in the nonlinear optimization.

We also obtained fitted values for the change in heat capacity on binding, ΔC_P, either by taking the slope of the direct binding enthalpies (direct method), or by taking the average of the fitted ΔC_P over five temperatures from the van't Hoff method (described below). This allowed an additional comparison of direct and van't Hoff method calculations.

2.5. Van't Hoff method

The van't Hoff enthalpies, $\Delta H_{VH,}$ were computed by applying the van't Hoff relationship to a series of binding free energies, ΔG(T), at several different temperatures, T. We used an integrated form of the van't Hoff relationship, which is sometimes termed the Gibbs-Helmholtz equation [17]:

$$\Delta G^o(T_j) \approx \Delta H(T_i) - T\Delta S^o(T_i) + \Delta C_P(T_i)\left[(T_j - T_i) - T_j \ln\left(\frac{T_j}{T_i}\right)\right] \quad (5)$$

Here, T_i and T_j are two temperatures, ΔG^o and ΔG^o are the standard binding free energy and entropy at standard concentration (1 M), ΔC_P is the change in heat capacity on binding, and the thermodynamic quantities are written as explicit functions of temperature. Note that the approximation in Eq. (5) becomes an equality when the heat capacity is independent of temperature. For each host-guest pair, ITC experiments were run at five temperatures, T_i, i = 1,2,3,4,5, and the binding free energy at each temperature was computed as $\Delta G^o(T_j) = -RT\ln K(T_j)$, where $K(T_j)$ is the measured binding constant. The van't Hoff enthalpy at temperature T_i, $\Delta H_{VH}(T_i)$, was then assigned by using nonlinear optimization to adjust the three quantities $\Delta H_{VH}(T_i)$, $\Delta S^o(T_i)$ and $\Delta C_p(T_i)$ so that the values of $\Delta G^o(T_j)$ computed by Eq. (5) across all five temperatures were best fit to their corresponding experimental values. Weighted, non-linear fitting was used to account for the uncertainties in $\Delta G^o(T_i)$ in obtaining the van't Hoff enthalpies. We found that setting each value of T_i as a reference temperature to compute $\Delta H_{VH}(T_i)$ in this way yielded better agreement with the direct enthalpies than an alternative procedure in which only one value of T_i was used as a reference temperature for all five values of $\Delta H_{VH}(T_i)$. In this manner, ΔC_p is fit at each temperature, with the reported ΔC_p being the average value across all temperatures. We also find that the most reliable van't Hoff enthalpies are typically obtained from temperature-dependent free energies computed with the stoichiometry, N, allowed to float (see Results), so this procedure is used for all van't Hoff enthalpies reported here, except as otherwise noted.

Experimental uncertainties in the van't Hoff enthalpies were obtained by a bootstrapping procedure similar to that used for the direct enthalpies. For each set of bootstrapped Wiseman plots (five Wiseman plots corresponding to a given concentration error), the fitted $\Delta G^o(T_j)$ values from each plot were used to compute $\Delta H_{VH}(T_i)$. Thus, for each

set of Wiseman plots a corresponding set of $\Delta H_{VH}(T_i)$ values is generated, and the resultant distribution of these is used to determine the mean and standard deviation.

2.6. Statistical measures

We wish to quantify the deviations between direct and van't Hoff enthalpies under various conditions and with various assumptions regarding experimental error (see Results). To do this, we use procedures described above to generate N_{sample} artificial sample datasets with the desired levels of error, where each dataset includes results at $N_{temp} = 5$ temperatures (see above), and to use nonlinear fitting to derive values of direct and van't Hoff enthalpy from these artificial data. We then define the discrepancy, D, between the direct and van't Hoff enthalpies as their mean unsigned difference:

$$D = \frac{1}{N_{temp}} \frac{1}{N_{samp}} \sum_{i=1}^{N_{temp}} \sum_{j=1}^{N_{samp}} |\Delta H_{VH,ij} - \Delta H_{Direct,ij}| \quad (6)$$

Except as otherwise noted, we allow the stoichiometry, N, to float when computing values of ΔG_o for use in obtaining the van't Hoff enthalpies (see above). However, we consider allowing N to float or be held fixed at 1 when computing the direct enthalpies, and examined how this choice affected the discrepancy D. To quantify this effect, we define $\Delta D = D_{float} - D_{fix}$, where the two values of D are computed with values of ΔH_{Direct} obtained with N allowed to float and N held fixed, respectively.

We similarly quantified the error in direct enthalpies computed by fitting to model calorimetry data with experimental error, relative to the nominal or "true" error-free values used to generate the data, as their mean unsigned error:

$$E = \frac{1}{N_{temp}} \frac{1}{N_{samp}} \sum_{i=1}^{N_{temp}} \sum_{j=1}^{N_{samp}} |\Delta H_{True,i} - \Delta H_{Direct,ij}| \quad (7)$$

We also analogously defined $\Delta E = E_{float} - E_{fix}$, which is positive if computing the direct enthalpies with N allowed to float leads to higher errors than treating N as fixed, and is negative in the opposite case.

2.7. Code availability

The code used to generate and analyze model data is available at (https://github.com/GilsonLabUCSD/Isothermal-Titration-Calorimetry). It allows one to generate model ITC data, based on assumed binding free energies and enthalpies over a temperature range, apply estimated concentration error and heat noise, use statistical sampling to generate model enthalpograms, and use these to obtain values of the direct and van't Hoff enthalpies. It can thus be used to estimate the level of van't Hoff consistency one should expect for a given binding system and estimated levels of heat and concentration error.

3. Results

We first report experimental ITC measurements of the binding thermodynamics of two adamantyl based guests to beta-cyclodextrin at multiple temperatures. Good agreement is obtained between the direct (ΔH_{Direct}) and van't Hoff (ΔH_{VH}) binding enthalpies obtained from the same data. We then use simulated data to investigate the consequence of two key sources of experimental error, heat error and concentration error [16], for the consistency between ΔH_{Direct} and ΔH_{VH}. We examine the level of consistency expected under typical experimental conditions, and provide evidence that lack of consistency can serve as an indicator of high experimental error.

3.1. Experimental van't Hoff consistency

To ascertain the level of van't Hoff consistency that could be observed experimentally with a modern calorimeter and appropriate experimental design, we used ITC to study the binding thermodynamics of two host-guest systems, beta-cyclodextrin with amantadine and rimantadine. These experiments avoid potential technical pitfalls, because all reactants are reasonably soluble, and because beta-cyclodextrin, placed in the syringe, has a minimal heat of dilution (Supplemental Fig. 1) and carries no electric charge. It is worthwhile to note that even the ionic guests give negligible heats of dilution (Supplemental Fig. 1). It is reasonable to assume then, that most of the error in the measurements will stem from heat and concentration error, which we have characterized experimentally [16].

For both host-guest pairs, measurements were taken at five temperatures, ranging from 300 K to 340 K. The average free energies and direct enthalpies of binding over two sets of replicates for both guests are presented in Table 1, with errors assigned as described in Materials and methods. It is worth noting the clear changes in these enthalpograms as temperature increases (Fig. 1): reduced sharpness of the sigmoidal Wiseman plots indicate drops in affinity, and, for rimantadine, the injection peaks are clearly taller at high temperature, indicating a more favorable binding enthalpy. This combination of changes implies entropy-enthalpy compensation, as borne out by the fitted data in Table 1. Note that these trends in free energy and enthalpy with T cannot be attributed to concentration error, because the same solutions were used across the range of temperatures in all cases. It is also worth noting that the results for rimantadine at 300 K agree with previously published binding measurements for this host-guest pair to within 0.1 kcal/mol [18].

The direct enthalpies are within 95% confidence intervals (about two standard deviations) of van't Hoff enthalpies obtained from the same data (Table 2, Fig. 2), with mean unsigned deviation between van't Hoff and calorimetric enthalpies of D = 0.16 kcal/mol and D = 0.39 kcal/mol for amantadine and rimantadine, respectively. However, the slopes of the van't Hoff enthalpies with T deviate somewhat from those of the direct enthalpies. This pattern is suggestive of heat error somewhat in excess of our estimates, rather than concentration error, as indicated in Fig. 4, panels b,d, and associated text (below). The heat capacity changes, ΔC_p, derived from the two methods also agree reasonably well; the values for amantadine are 74.6 and 70.1 cal/mol/K, for direct and van't Hoff, respectively; for rimantadine, the corresponding results are 72.7 and 93.9 cal/mol/K. The heat capacities come from the second derivative of the binding free energies with respect to temperature, and thus are more sensitive to noise than the enthalpies. Additional measurements performed with reduced concentrations of both host and guest (Supplementary Fig. 3) are consistent with the present results. This indicates that the binding thermodynamics are independent of concentration in this range, and thus that the solutions used are close to ideal.

3.2. Propagation of experimental error to calorimetric and van't Hoff enthalpies

Here, we use simulated data to characterize the distinct ways in which concentration error and heat error propagate to ΔH_{Direct} and ΔH_{VH}, to test whether severe inconsistency between these two quantities could be useful to flag data with high errors, and to estimate the level of agreement between these quantities to be expected under typical conditions for near-ideal solutions. Ideal Wiseman plots at five temperatures were generated for values of K, ΔH^o, and ΔC_p similar to those of rimantadine binding beta-cyclodextrin (Supplementary Table 1), and assuming 1:1 binding (N = 1) (Fig. 3). As a methodological check, we confirmed that fitting to these ideal Wiseman plots data yielded values of ΔH_{Direct} and ΔH_{VH} that agree essentially perfectly with each other and with the binding enthalpy used to generate the model

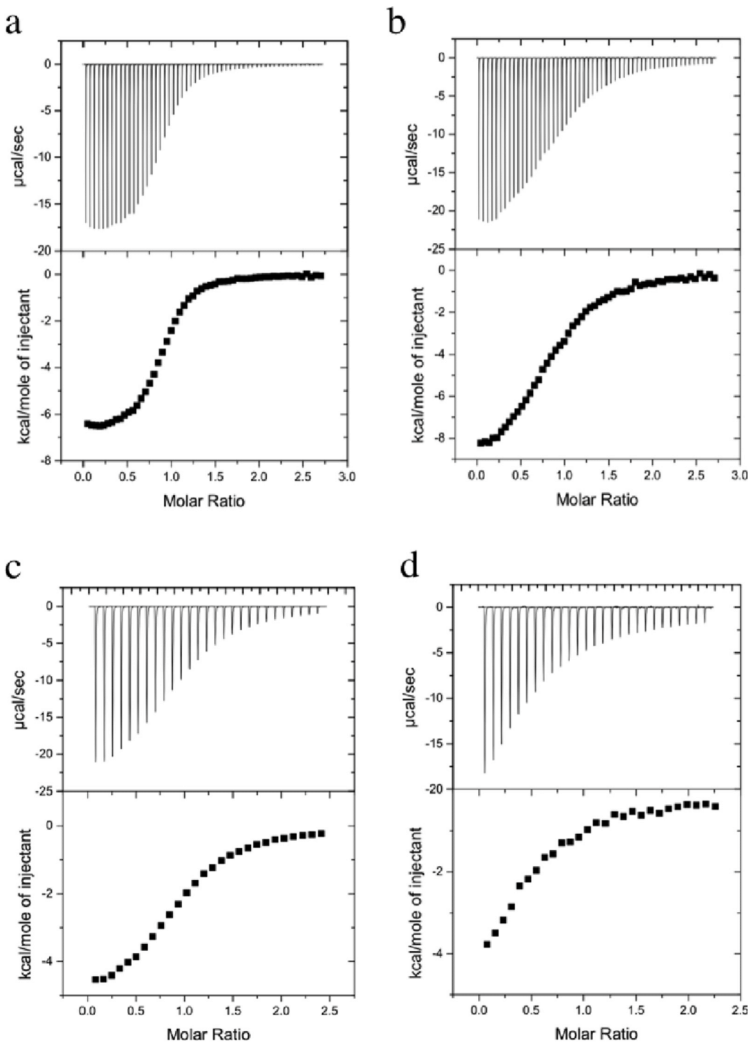


Fig. 1. Sample experimental plots for beta-cyclodextrin binding rimantadine (a,b) and amantadine (c,d). Panels a and c show measurements at 300 K, panels b and d show measurements at 340 K. For each panel, the top half corresponds to the raw enthalpogram and the bottom to integrated enthalpograms (Wiseman plots). For each guest compound, the injectant concentrations and volumes are the same at 300 K and 340 K, as are the cell concentration. The ITC C values for these experiments are as follows: 50 (a), 6 (b), 10 (c), 3 (d).

data (data not shown). We then added heat error or concentration error to each ideal Wiseman plot. In order to simulate a series of ITC measurements using the same solutions across different temperatures, the same erroneous concentrations were used across each temperature series, though each temperature series used a different assumed concentration error. Bootstrap resampling was used to generate 1000 such independent temperature series, resulting in 1000 values of ΔH_{Direct} and ΔG^o for each temperature point. The values of ΔG^o were furthermore used to fit 1000 van't Hoff enthalpies at each temperature, based on Eq. (5). The magnitude and character of both the concentration error and heat error used here are chosen to reflect typical experimental conditions, as determined in our prior empirical study [16].

A central observation of this work is that realistic levels of concentration error and heat error lead to strikingly different patterns of error in ΔH_{Direct} and ΔH_{VH}, as illustrated by representative data drawn from the 1000 bootstrapped datasets for N floating (Fig. 4) and N fixed at 1 (Fig. 5). With N floating, pure concentration error propagates significantly to ΔH_{Direct} (Fig. 4a), but negligibly to ΔH_{VH} (Fig. 4b). That pure concentration error has minimal effect on the van't Hoff enthalpies presumably derives from our assumption that the same solutions are used across all temperatures, as this leads to a small and nearly constant shift in ΔG^o across T, which has little effect on the slope and hence on ΔH_{VH}. Conversely, with N floating, heat error propagates only slightly to ΔH_{Direct} (Fig. 4c), but propagates strongly to ΔH_{VH} (Fig. 4d), leading to plots of van't Hoff enthalpy

Table 1
Binding enthalpies and free energies (kcal/mol) for amantadine and rimantadine with beta-cyclodextrin, from 300 to 340 K. Uncertainties are one standard deviation, generated from bootstrapping during fitting as described in Materials and methods. For amantadine, the C values range from 2 to 10 over the temperatures series; for rimantadine, the C values range from 10 to 50.

	Amantadine			Rimantadine		
	ΔH	ΔG°	TΔS	ΔH	ΔG°	TΔS
300 K	− 5.29 ± 0.04	− 5.29 ± 0.01	0.00 ± 0.04	− 6.75 ± 0.04	− 6.23 ± 0.01	0.52 ± 0.04
310 K	− 6.08 ± 0.06	− 5.27 ± 0.01	0.81 ± 0.06	− 7.56 ± 0.04	− 6.20 ± 0.01	1.36 ± 0.04
320 K	− 6.55 ± 0.07	− 5.27 ± 0.01	1.28 ± 0.07	− 8.34 ± 0.05	− 6.19 ± 0.01	2.15 ± 0.05
330 K	− 7.42 ± 0.10	− 5.18 ± 0.01	2.24 ± 0.10	− 8.98 ± 0.06	− 6.11 ± 0.01	2.87 ± 0.06
340 K	− 8.13 ± 0.41	− 5.16 ± 0.05	2.97 ± 0.43	− 9.67 ± 0.21	− 6.01 ± 0.02	3.66 ± 0.23

Table 2
Direct and van't Hoff binding enthalpies (kcal/mol) for amantadine and rimantadine binding beta-cyclodextrin from 300 to 340 K. Errors shown are one standard deviation, generated from bootstrapping, as outlined in the Materials and methods section.

	Amantadine		Rimantadine	
	ΔH_{Direct}	ΔH_{VH}	ΔH_{Direct}	ΔH_{VH}
300 K	− 5.29 ± 0.04	− 5.01 ± 0.58	− 6.75 ± 0.04	− 5.97 ± 0.43
310 K	− 6.08 ± 0.06	− 5.80 ± 0.25	− 7.56 ± 0.04	− 6.97 ± 0.22
320 K	− 6.55 ± 0.07	− 6.56 ± 0.27	− 8.34 ± 0.05	− 7.92 ± 0.15
330 K	− 7.42 ± 0.10	− 7.29 ± 0.59	− 8.98 ± 0.06	− 8.84 ± 0.33
340 K	− 8.13 ± 0.41	− 8.00 ± 0.93	− 9.67 ± 0.13	− 9.73 ± 0.53

versus temperature whose slopes deviate from their true value. It is worth noting that the realistic levels of concentration and heat error used here lead to errors in fitted ΔG° on the order of only hundredths of a kcal/mol (Supplementary Fig. 1a, b), smaller than those previously assumed in studying uncertainties in van't Hoff binding enthalpies [19].

Fixing N at its nominal value (N = 1 in the present study) can often improve the reliability of thermodynamic data from ITC measurements [16,20], so it is interesting to examine how fixing N influences error propagation into ΔH_{Direct} and ΔH_{VH}. We find that, with N fixed, concentration error has a somewhat reduced effect on ΔH_{Direct} (Fig. 5a vs. Fig. 4a), but it now propagates dramatically to ΔH_{VH} (Fig. 5b vs Fig. 4b). However, heat error propagates to both values about the same with N fixed (Fig. 5c,d) as with N floating (Fig. 4c,d).

Finally, we examined the levels of error expected when determining van't Hoff enthalpies under common calorimetric operating conditions, assuming the solutions are essentially ideal. This was done by using Monte Carlo sampling over heat errors and concentration errors, as previously described [16]. The resulting uncertainties in the enthalpies depend on the assumed uncertainty in concentrations, and on whether N was treated as fixed or allowed to float (Table 3), but for a typical 2% concentration uncertainty [16], the uncertainties in van't Hoff enthalpies run about 3–4%. This is consistent with the discrepancies D of ~2–4% in the present experimental data (above). This analysis raises the possibility that unexpectedly large discrepancies between van't Hoff and direct enthalpies could be used to flag potentially problematic data. This idea is considered in the following subsection.

3.3. Relationship between van't Hoff consistency and magnitude of error

Given that ΔH_{Direct} and ΔH_{VH} agree perfectly with each other in the absence of experimental error, and that different sources of error result in distinct patterns of error in these two quantities, we conjectured that large inconsistencies between ΔH_{Direct} and ΔH_{VH} could be utilized to flag datasets that contain relatively high levels of error of any sort. Here we test this idea using simulated data. For these model measurements, we consistently allowed N to float in computing ΔH_{VH}, because this markedly reduced sensitivity to concentration error, as noted above.

Focusing first on pure concentration error (i.e., without heat error), we find that increasing the assumed level of concentration error from 2% to 10% leads to a simultaneous increase in *E*, the deviation of ΔH_{Direct} from the true value used to generate the model data, and *D*, the discrepancy between ΔH_{Direct} and ΔH_{VH}, whether N is floated or held fixed at 1 in computing ΔH_{Direct} (Fig. 6) (see Eqs. (6), (7) for definitions of *D* and *E*, respectively). The same trend is observed when a realistic level of heat error is added, though only when concentration is error rises above 4% (Fig. 6 c,d). Thus, for both pure concentration error and for concentration error combined with heat error, a dataset which gives

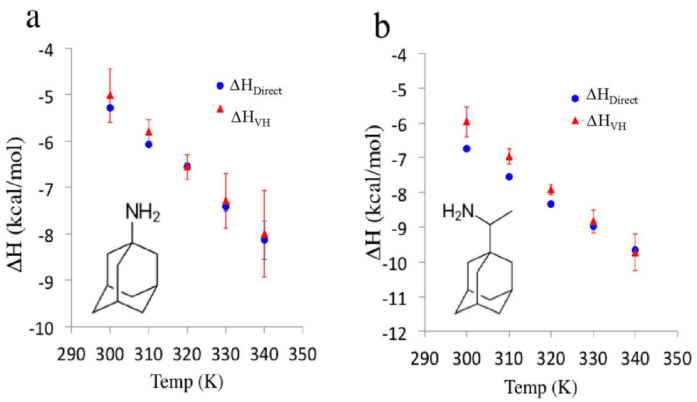


Fig. 2. Measured direct and van't Hoff binding enthalpies versus temperature. a) beta-cyclodextrin and amantadine. b) beta-cyclodextrin and rimantadine. Error bars shown are one standard deviation, generated from bootstrapping, as outlined in the Materials and methods section.

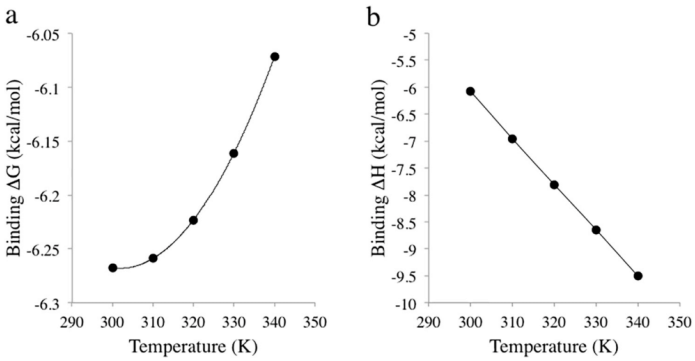


Fig. 3. Ideal binding free energies and enthalpies, corresponding to ideal Wiseman plots. (A) Binding free energies from ideal Wiseman plots at five temperature points (B) Binding enthalpies from ideal Wiseman plots at five temperature points.

good agreement between ΔH_{Direct} and ΔH_{VH} also tends to give a more accurate value of ΔH_{Direct}.

Results analogous to those in Fig. 6 but with concentration error in only the syringe or cell solutions are provided in Supplementary Figs. 2 and 3. These added results provide insight into how floating versus fixing N affects the relationship between the metrics *D* and *E*, through the differential effect of the treatment of N upon the propagation of cell and syringe concentration errors [16,20].

The results above show that agreement between direct and van't Hoff enthalpies is an indicator of data quality. Here, we extend this idea by inquiring whether this level of agreement can be used to decide whether it is better to keep N fixed or allow it to float when extracting direct binding enthalpies from calorimetric data. That is, it may be possible to discern whether floating or fixing N will decrease error in this fitted enthalpy. We used the model system used above to generate 1000 sets of five temperature points, and compared the van't Hoff discrepancies, *D*, and calorimetric

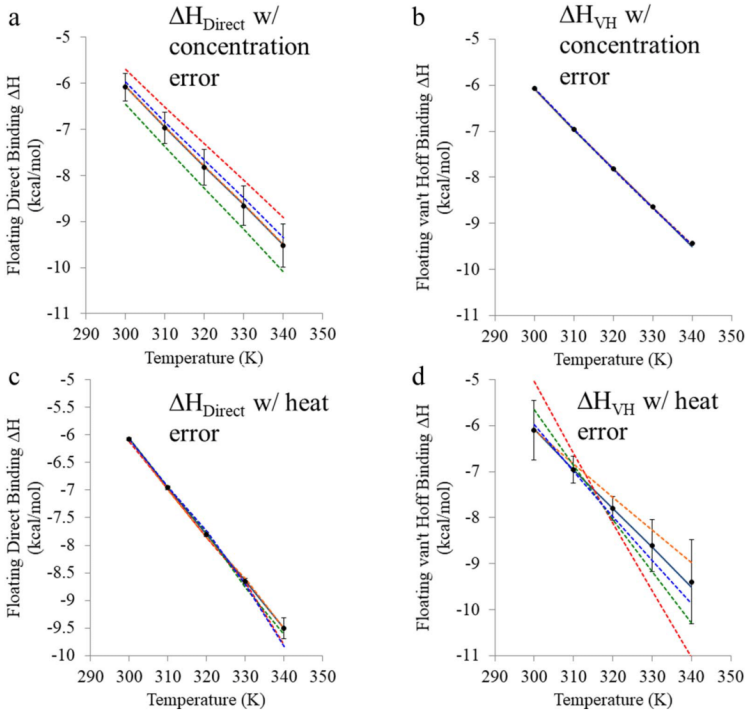


Fig. 4. Effect of heat and concentration error on direct and van't Hoff binding enthalpies based on binding free energies fitted with stoichiometry parameter N floating. The solid black line in each panel represents the nominal binding enthalpies used to generate the fitted Wiseman plots, while each of the dashed, colored lines plot the binding enthalpies from one set of replicates; five examples of these are included in each panel. Error bars represent one standard deviation over 1000 replicates. For concentration error, 5% syringe and cell error was applied. For heat error, for 300–330 K points, 1% heat error and 0.13 μcal irreducible baseline error was added; for 340 K points, 3% heat error and 0.39 μcal irreducible baseline error was added. (A) Fitted direct binding enthalpy at each temperature point with concentration error applied (B) Fitting van't Hoff binding enthalpies at each temperature point concentration error applied (C) Fitted direct binding enthalpies at each temperature point with heat error applied. (D) Fitted van't Hoff binding enthalpies at each temperature with heat error applied.

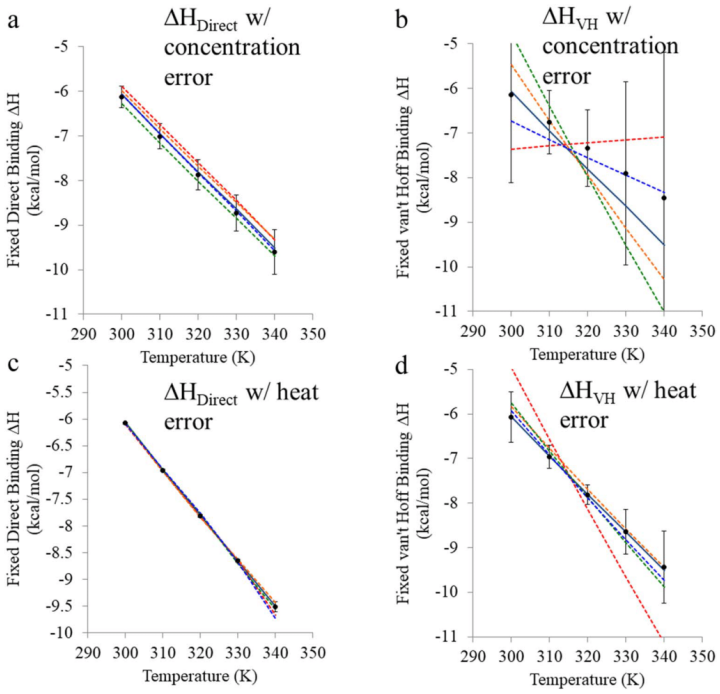


Fig. 5. Effect of heat and concentration error on direct and van't Hoff binding enthalpies with stoichiometry parameter N fixed at one. Dashed color lines indicate the binding enthalpies from one set of replicates; the solid line represents the nominal binding enthalpies used to generate the fitted Wiseman plots. Error bars represent one standard deviation over 1000 replicates. For concentration error, 5% syringe and cell error was applied. For heat error, for 300–330 K points, 1% heat error and 0.13 μcal irreducible baseline error was added; for 340 K points, 3% heat error and 0.39 μcal irreducible baseline error was added. (A) Fitted direct binding enthalpy at each temperature point with concentration error applied (B) Fitting van't Hoff binding enthalpies at each temperature point concentration error applied (C) Fitted direct binding enthalpies at each temperature point with heat error applied. (D) Fitted van't Hoff binding enthalpies at each temperature with heat error applied.

enthalpy errors, *E*, when N was floated and fixed. We then plotted $\Delta E = E_{float} - E_{fixed}$ against $\Delta D = D_{float} - D_{fixed}$ to look for correlation (Fig. 7). Note that the van't Hoff enthalpies here are computed based on binding free energies obtained with N floating; only the direct enthalpies were computed with either N fixed or N floated.

When zero heat error is assumed, the deviation of the direct enthalpies from the true enthalpy correlates perfectly with the discrepancy between the van't Hoff and direct enthalpies (Fig. 7a). Indeed, a linear regression of ΔE against ΔD gives a slope and correlation coefficient of essentially unity, and a y-intercept of essentially zero. In this idealized situation, then, one may determine whether to float or fix N based on which fitting method gives better agreement between the van't Hoff and direct enthalpies (lower D). Overall, heat error (at the level estimated for the VP-ITC instrument [16]) adds scatter to the plot (Fig. 7b), but the relationship between the discrepancy and error remains, with a still high R^2. This relationship allows for a straightforward "consistency check" for measurements done over several temperatures, and may be useful in choosing whether to treat N as fixed or allow it to float, for a given dataset.

4. Discussion

The present study addresses a long-standing and continuing discussion in the calorimetry literature as to whether one should expect agreement between the binding enthalpy obtained directly from an ITC experiment and the binding enthalpy obtained by the van't Hoff method. We build on prior observations of error propagation in ITC [16,20–22] and expand this to include enthalpies obtained through the application of the van't Hoff equation to binding free energies obtained with ITC. We determine the level of agreement between van't Hoff and direct enthalpies to be expected under typical operating conditions, in the absence of significant solution non-ideality, and provide new measurements that showing agreement well within the expected uncertainties. We conjecture that the quality of agreement, compared with some prior studies, traces to two factors. The first is our use of more modern, lower-noise, instrumentation than that used in some of the older studies. The second is our use of an experimental design that avoids injecting a charged solute into the ITC cell, a practice which we suspect could generate complicating non-idealities.

Furthermore, we have shown, using model binding data, that key sources of experimental error propagate differently to the direct and the direct and van't Hoff enthalpies, leading to discrepancies between these two fitted quantities. As a consequence, inconsistency between the direct and van't Hoff enthalpies greater than those expected for typical operating conditions may be a useful flag for problematic data. That is, the consistency between these quantities correlates with the accuracy of the results. Similarly, one may use the level of agreement between van't Hoff and direct enthalpies to decide whether it is preferable to treat N as fixed or to float it. The broad picture that agreement between van't Hoff and direct enthalpy correlates with accuracy makes sense intuitively, given that, as argued below, the direct and van't Hoff results should in principle agree to within modest uncertainties attributable to heat and concentration error. Accordingly, excessive deviations between them must be attributable to some additional error in the experiment.

The following subsections consider in more detail prior reports of large discrepancies between van't Hoff and direct calorimetric enthalpies, consider the role of experimental error as the cause of such discrepancies, and examine prior theoretical arguments that the direct and van't Hoff enthalpies should not in fact be expected to agree. It

Table 3
Errors in calorimetric (Error_direct) and van't Hoff (Error_VH) enthalpies in modeled data, expressed as mean unsigned error. Each row corresponds to a set of model data with the specified error applied, with the data being fit allowing N to float or N fixed, as described in the Materials and methods. Heat errors, considered in the lower half of the table, include 1% proportional heat error at 300 K and 3% heat error at 340 K, as previously determined [16]. The modeled data uses the same C values and input parameters as described in the Materials and methods section.

	Syr Err	Cell Err	Error_direct (%MUE)	Error_VH (%MUE)
Without heat error				
Float	2%	2%	1.6	0.1
	4%	4%	3.4	0.1
	6%	6%	4.4	0.2
	8%	8%	6.5	0.2
Fixed	2%	2%	1.4	1.1
	4%	4%	2.6	3.0
	6%	6%	4.1	4.5
	8%	8%	6.9	11.9
With heat error				
Float	2%	2%	1.5	3.8
	4%	4%	3.1	4.1
	6%	6%	4.5	4.1
	8%	8%	7.1	4.1
Fixed	2%	2%	1.2	3.4
	4%	4%	2.9	4.5
	6%	6%	4.4	6.4
	8%	8%	4.8	10.7

should be noted that the scope of this study only concerns comparison between van't Hoff and direct enthalpies obtained from isothermal titration calorimetry. We do not consider other possible errors in van't Hoff enthalpies obtained from other experimental methods of measuring binding affinities, such as spectrophotometry, NMR, and surface plasmon resonance.

4.1. Past discrepancies between direct and van't Hoff enthalpies

In contrast with the present study, a number of prior contributions, including those of Sturtevant [4,5,23], Chaires [19], and Eggers [12], have noted persistent discrepancies between van't Hoff and calorimetric enthalpies and argued that the van't Hoff equation may be flawed, or that there are other factors in play that can make this quantity impossible to obtain accurately. One reason for our different experience may be that we chose reagents which we expected would have activity coefficients near unity at the experimental concentrations, due to their low charge densities, so that our experiments would provide valid equilibrium constants. Thus, cyclodextrin is a large, neutral molecule, and the amantadine and rimantadine guests combine a bulky hydrophobic moiety with a singly charged ammonium group. That their activity coefficients were in fact close to unity is evidenced by the very low heats of dilution of the reagents (Supplementary Fig. 1), and the fact that halving concentrations produced no significant change in the fitted affinity and binding enthalpy (Supplementary Fig. 2). We suspect that strong nonideality was a major contributor to the van't Hoff discrepancies seen in Eggers and Sturtevant's studies, as they used reagents with much higher charge densities,

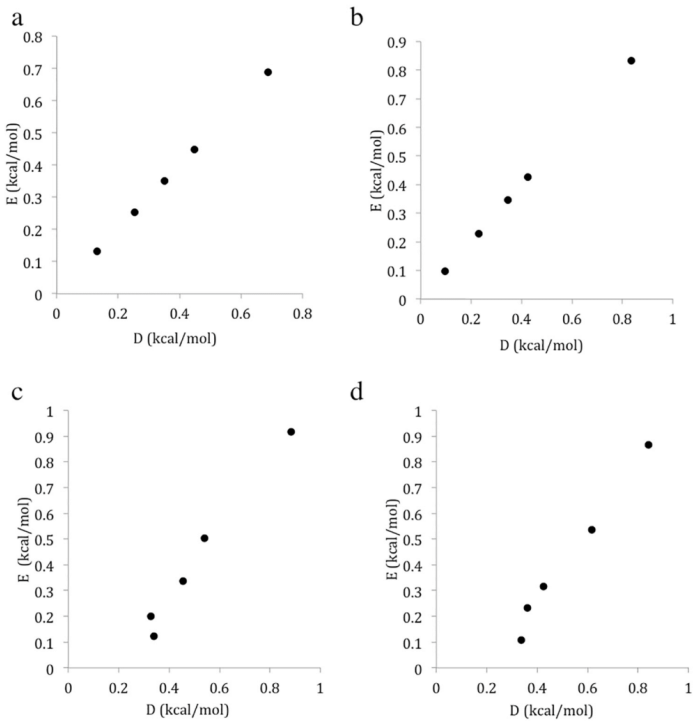


Fig. 6. Correlation between *D*, the discrepancy between direct and van't Hoff enthalpies and *E*, the error in fitted direct enthalpies. As points progress from bottom to top, they represent 2, 4, 6, 8, and 10% concentration error added to both syringe and cell. (A) Correlation between direct enthalpy error and discrepancy while N is allowed to float, with concentration error only. (B) Correlation between direct enthalpy error and discrepancy while N is fixed at one, with concentration error only. (C) Correlation between direct enthalpy error and discrepancy while N is allowed to float, with concentration error and 1% heat error added. (D) Correlation between direct enthalpy error and discrepancy while N is fixed at one, with concentration error and 1% heat error.

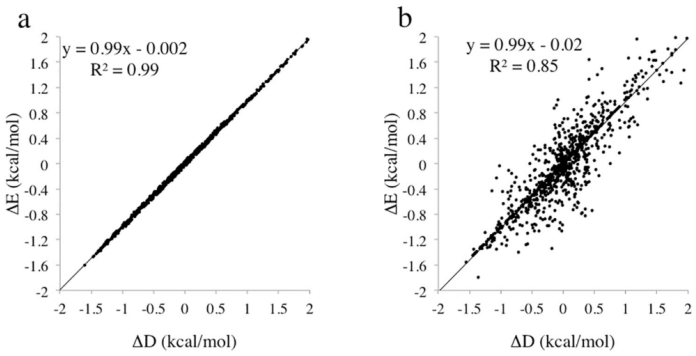


Fig. 7. Correlation between difference in float and fix error in fitted ΔH and difference in float and fix van't Hoff discrepancy in modeled data. Each point corresponds to a set of model ITC measurements at temperatures 300, 310, 320. 330, and 340 K, with up to 10% concentration error added to both cell and syringe and either no heat error (panel A) or with heat error based on our own experimental measurements (panel B). Here, ΔD represents the mean deviation between van't Hoff and directly fitted enthalpies across all temperatures (Eq. (6)), and ΔE represents the average error in the fitted ΔH across all temperatures (Eq. (7)).

including divalent cations and EDTA, which has four carboxylic acid groups. Binding measurements for highly nonideal solutions do not yield valid equilibrium constants, unless steps are taken to explicitly correct for the nonideality, and using invalid equilibrium constants in the van't Hoff equation will yield incorrect estimates of the binding enthalpy. Although Sturtevant argued that nonideality should not be an issue in his experiments, this was not experimentally documented. Moreover, his measurements for a less charged pair of reagents, cyclodextrin with heptanoate, provided “puzzling” good van't Hoff consistency [5]. Eggers' study almost certainly was carried out under nonideal conditions, as he observed large, temperature-dependent, variations in apparent binding affinity as a function of concentration. Applying the van't Hoff equation to these uncorrected apparent quantities could have yielded erroneous enthalpies.

Another reason for the relatively good agreement between van't Hoff and direct enthalpies observed here is our use of an instrument that has lower heat noise than those used in some of the older studies. For example, the VP-ITC is an improvement over the Calorimetry Sciences Corporation model 4200 used by Horn et al. (JR Horn, personal communication), and indeed, they noted that the van't Hoff discrepancies they obtained could be fully explained by instrument uncertainty [11]. Additionally, free energies measured with the VP-ITC under typical conditions have lower uncertainty than that assumed by Chaires in his modeling study of van't Hoff discrepancies [19]. Finally, it is worth remarking that Tellinghuisen recently observed significant differences between calorimetric and van't Hoff enthalpies, using a modern instrument, but did not argue against the validity of the van't Hoff equation and instead speculated that the discrepancy resulted merely from problems with handling the heats of dilution [6].

Potential explanations for apparent van't Hoff inconsistencies are considered further in the following subsections.

4.2. Sources of van't Hoff inconsistency: experimental error

Experimental or procedural errors are sometimes invoked to explain van't Hoff inconsistencies. Notably, a study by Chaires found that a small curvature in the curve of binding free energy versus temperature (i.e., a small value of ΔC_p) could be obscured by modest noise in ΔG, leading to highly inaccurate van't Hoff enthalpies. We find that this is mostly true: noise in ΔG stemming from heat error can significantly increase the scatter of fitted van't Hoff enthalpies, as evidenced by the results from modeled Wiseman plots. This can result from even low levels of uncertainty in ΔG, on the order of 0.05 kcal/mol. However, given the modest levels of heat error in the VP-ITC instrument, obtaining van't Hoff enthalpies in good agreement with direct enthalpies should not be as difficult as previously predicted, even with smaller ΔC_p's. The value of ΔC_p used for the model data in Sections 3.2 and 3.3 was -85 cal/mol/K, and the ΔC_p for both experimental datasets (which gave reasonable consistency) were ~ -70 cal/mol/K, suggesting that, even if ΔC_p is relatively small (Chaires' study assumed ΔC_p of 200 cal/mol/K), good consistency should still be achievable, at least in a well-controlled experimental setting.

We also find that concentration error has minimal effect on van't Hoff enthalpies, in contrast to the known sensitivity of the direct calorimetric enthalpy to concentration error [16,20,22,24]. This is presumably because we use the same solutions across for the full temperature series of experiments, so that any concentration error is constant across the series. Thus, any concentration error that is present shifts the apparent binding affinity at each temperature point by about the same amount, allowing the slope and curvature of the free energy versus temperature curve to remain essentially unaffected, and hence preserving the van't Hoff enthalpies. In contrast, if different solutions were used at the different temperatures, we expect that van't Hoff enthalpies would be greatly affected by concentration error in much the same way that they are affected by heat error, where the free energy at each temperature has independent scatter that affects the curvature and thus the binding free energies. The net effect would be a marked increase in the discrepancy between direct and van't Hoff enthalpies. It is therefore essential that measurements of van't Hoff enthalpies utilize the exact same solutions for the measurements at all temperatures. We guess that previous studies comparing direct and van't Hoff enthalpies used the same solutions across all measurements, but this methodological detail is rarely specified, so the use of different solutions at the different temperatures might, in some cases be an added source of discrepancy.

As noted previously [6,21], the van't Hoff equation applies to standard thermodynamic quantities, which pertain to ideal solutions, so apparent deviations between direct and van't Hoff binding enthalpies can arise if the solutions used deviate significantly from ideality, unless pains are taken to correct back to standard state. Carrying out such corrections would appear to require knowledge of the activity coefficients of the reactants and product, both alone and as mixtures, and perhaps also knowledge of the dependency of the partial molar enthalpies of these species upon concentration. These quantities are often not available, so corrections cannot easily be made. However, as previously emphasized [21], one may at least determine whether the binding free energy and enthalpy are insensitive to concentration at the desired experimental conditions, as done in the present study.

In the present study, then, the greatest contribution to inconsistency between van't Hoff and direct enthalpies seems to be the heat error intrinsic to isothermal titration calorimetry instruments. Despite the inevitable presence of these errors, given a robust C value across the temperature range (such as the range of 5 to 50 used in the present models), and based on current estimates of heat error [16], we

anticipate that consistency should be achievable to within about 4–5% mean unsigned error between van't Hoff and calorimetric enthalpies. Given that earlier ITC instruments had larger heat errors than the instrument used here, some past reports of van't Hoff inconsistency likely stemmed from this class of experimental error.

4.3. Sources of van't Hoff inconsistency: theoretical

Theoretical explanations have also been proposed to explain apparent inconsistencies between direct and van't Hoff binding enthalpies. Thus, Weber argued that the van't Hoff equation is fundamentally wrong, due to confusion between the entropy change of the reactants themselves and the entropy change of the whole system [7,8], but this line of reasoning was convincingly rebutted [9,10] and is no longer current in the literature. Another theoretical argument came from Sturtevant, who suggested that solvent effects play a role in discrepancies between van't Hoff and calorimetric enthalpies [4,5]. In this view, the rearrangement of solvent around the solutes on binding leads to an additional free energy term not accounted for when calculating $d\Delta G^o/dT$ to generate van't Hoff enthalpies. Castellano and Eggers [12] have recently elaborated this idea, arguing that the standard binding free energy comprises not only the customary difference in standard chemical potentials of the reacting solutes (e.g. a protein, a ligand, and their complex), but also an added contribution, termed the desolvation free energy, from the change in chemical potential of solvent molecules released from the binding surfaces to bulk upon binding. They suggest that this term is missing from conventional treatments of the thermodynamics of reactions in solution, and that this concept can be used to explain many experimental observations of discrepant direct and van't Hoff enthalpies.

We disagree with this view, and now present a brief statistical thermodynamic derivation of the van't Hoff equation, which shows that this relation holds true even when solvent is fully accounted for, and that changes in solvation of the reactants on binding do not require any modification to solution binding theory. We note that, although the van't Hoff equation can be derived based purely on thermodynamics the statistical thermodynamics perspective is informative here because it makes the roles of the various molecular species explicit.

The standard free energy of reaction is the difference between the standard chemical potentials of the product and reactant species [25], which here are the bound and free species [26]. This is because the standard chemical potential of a molecular species is the change in free energy of the system when one mole of the species is added to the system. When the reaction proceeds forward by one mole under standard conditions, one mole of each reactant is consumed and one mole of product is generated. Consequently, the change in the free energy of the system is the difference in chemical potentials:

$$\Delta G^o = -RT \ln K = \mu^o_{AB} - \mu^o_A - \mu^o_B \tag{8}$$

Here $\mu_X{}^o$is the standard chemical potential of species X = AB, A or B, where A and B bind to form the complex AB. Note that the standard chemical potentials pertain to hypothetical ideal solutions at standard concentration (typically 1 M), and that K is a concentration-independent quantity, by definition. Because no water is added or consumed in the course of the reaction, the chemical potential of water does not enter explicitly. Instead, the presence of water enters through its influence on the chemical potentials of A, B, and AB. Thus, the standard chemical potential of species X is given by [14,26]

$$\mu^o_X = -RT \ln\left(\frac{8\pi^2}{C^o}\frac{\int e^{-\frac{U(r_X,r_{solv})}{RT}} dr_X dr_{solv}}{\int e^{-\frac{U(r_{solv})}{RT}} dr_{solv}}\right) \tag{9}$$

Here $8\pi^2$ accounts for the free rotation of the solute, C^o is the standard concentration (typically 1 M), r_x represents the internal coordinates of species X, r_{solv} represents the coordinates of the water and other solvent molecules in the coordinate frame of the solute, $U(r_X,r_{solv})$ represents the potential energy of the solute-solvent system as a function of these coordinates and $U(r_{solv})$ is the potential energy of the solvent in the absence of the solute. (See refs 17 and 18 for details.) This expression accounts for all favorable and unfavorable interactions of water with the solute, so the difference in chemical potentials in ΔG^o fully accounts for the fact that the bound complex has less solvent-exposed surface than the unbound reactants. Now, recognizing that

$$\frac{\partial}{\partial T}\left(\ln \int e^{-U(r)/RT} dr\right) = \frac{1}{RT^2}\langle U\rangle \tag{10}$$

and using the equations above, it is straightforward to complete a derivation of the van't Hoff equation [14]:

$$\frac{\partial}{\partial T}\ln K = \frac{(\langle U_{AB}\rangle - \langle U_{solv}\rangle) - (\langle U_A\rangle - \langle U_{solv}\rangle) - (\langle U_B\rangle - \langle U_{solv}\rangle)}{RT^2} = \frac{h_{AB} - h_A - h_B}{RT^2} = \frac{\Delta H^o}{RT^2} \tag{11}$$

Here, we have used the fact that h_x, the standard partial molar enthalpy of solute X, is the change in mean energy when a mole of solute is added to pure solvent, with the simplifying approximation that any net volume change is negligible. Thus, the van't Hoff relation follows directly from the statistical thermodynamics of ideal solutions, and it fully accounts for the changes in solute-water interactions that occur on binding.

We therefore see no basis for the view of Castellano and Eggers that changes in reactant solvation on binding can lead to violations of the van't Hoff equation. Furthermore, the chemical potential of water, which is defined as the change in the free energy of the system when another mole of water is added, is not a function of position. Therefore, we do not believe there is a sound theoretical basis for their claim that water at the solute interface has a different chemical potential from bulk water in the same flask. Analogous reasoning holds for other components of the system, such as protons. Thus, Horn et al. showed that linked equilibria should not affect van't Hoff enthalpies, and therefore can be discounted as a possible explanation for discrepancies between van't Hoff and direct enthalpies [27].

If the reworking of solution thermodynamics proposed by Castellano and Eggers does not hold, then another explanation must be sought for the apparently severe discrepancy between van't Hoff and direct enthalpies in their EDTA-Ca ITC experiments. They observe good agreement between the direct and van't Hoff enthalpies in experiments where the reactants are at low concentration, but increasing inconsistency with rising concentrations. This pattern suggests a role for non-ideality, which also tends to become more marked with increasing concentration. Non-ideality can play a role because the equilibrium constant which enters the van't Hoff equation is, by definition, independent of concentration, whereas Castellano and Eggers may instead have used an apparent equilibrium constant, by which we mean the equilibrium ratio of concentrations. When solutions are not ideal, this is a concentration-dependent quantity, and there is no reason to expect the van't Hoff equation to work if it is used in place of the equilibrium constant itself. In the present instance, we anticipate that the highly charged reactants, Ca^{2+} and $EDTA^{2-}$ or $EDTA^{3-}$, will have substantially lower activity coefficients at high concentration than at low concentration. However, the calcium-EDTA complex will likely have an activity coefficient near unity, because its net charge is near zero. Thus, raising the concentrations of all species will lead to an apparent weakening of binding, as was observed experimentally. For example, if at high concentration the activity coefficients of the reactants are $\gamma_C = \gamma_E = 0.5$, while that of the complex is $\gamma_{C.E} = 0.9$, then, given the usual expression for the equilibrium constant,

$$K = \frac{\gamma_{C.E} C_{C.E}}{\gamma_C C_C \ \gamma_E C_E} \tag{12}$$

the apparent equilibrium constant will be

$$K_{app} = \frac{C_{C.E}}{C_C C_E} = K\frac{\gamma_C \gamma_E}{\gamma_{C.E}} = 0.28K \quad (13)$$

Thus, the reduced activity of the reactants vs. the products, at high concentration, leads to less binding. Although one may insert K_{app} into the van't Hoff equation in place of K and thus obtain an apparent van't Hoff enthalpy that differs from the direct binding enthalpy, this inconsistency is not in fact an example of van't Hoff inconsistency, because the van't Hoff equation by definition uses K, not K_{app}. Additionally, the partial molar enthalpy of a solute may change with concentration, which would be another source of discrepancy in observed calorimetric and van't Hoff enthalpies. It is also worth emphasizing that the large change in the apparent equilibrium constant with concentration observed in the EDTA-Ca experiments is itself direct evidence that the solutions are not behaving ideally at high concentration. Thus, nonideal solution behavior likely explains the large apparent van't Hoff inconsistency seen by Castellano and Eggers for their measurements at high concentration, and their results do not imply that the long-standing theory of solution thermodynamics needs to be revised. Nonideality might also help explain other cases of apparent van't Hoff inconsistency, as previously emphasized by Pethica [21], particularly in cases where one or more of the reactants are highly charged.

Given the apparent occurrence of significant nonideality in prior calorimetric studies of small molecules, it is of interest to consider whether nonideality also comes into play in biomolecular applications. We are not aware of many measurements of protein activity coefficients, but a study of the protein alpha-chymotrypsin reported activity coefficients within a few percent of 1 for concentrations up to about 1 μM, under various salt conditions [28]. Thus, applications of ITC that use protein concentrations in this range or lower probably avoid serious nonideality, at least for the protein solute. The concern is greater, however, for highly charged solutes, such as nucleic acids. The presence of nonideality may be checked by looking for concentration dependence of the results, or by comparing van't Hoff and direct enthalpies, because, as mentioned above, the concentration dependent apparent K and partial molar enthalpy would cause discrepancies between the two enthalpies.

4.4. Using van't Hoff consistency to check for error

Given that the direct and van't Hoff enthalpies should, in principle, be in accord, an apparent discrepancy between these two quantities beyond that expected based on known experimental uncertainties can be a useful flag for the presence of experimental errors of various types, such as errors in solution concentrations, unexpectedly high heat errors, or marked nonideality of the solutes. This concept may be particularly useful in settings where high accuracy is desired. In addition, the van't Hoff consistency check can be used to guide the decision of whether to fix or float the stoichiometry parameter, N, in fitting ITC data, as our model calculations show that greater consistency correlates with greater accuracy.

5. Conclusions

This study yields several key observations regarding the nature of enthalpies obtained from both calorimetry (direct) and van't Hoff methods using an isothermal titration calorimeter:

- The direct and van't Hoff enthalpies should in principle agree well
- Inconsistency between the van't Hoff and direct enthalpies can results from concentration error, heat error, or solution nonideality.
- Inconsistency between the van't Hoff and direct enthalpies is a sign of experimental error.
- Concentration error has little effect on van't Hoff enthalpies when N is floated, given that the same solution is used for experiments at all temperatures.
- Concentration error can propagate strongly to direct enthalpies, with cell error propagating when N is fixed, and syringe error propagating when N is floated.
- Heat error propagates strongly to van't Hoff enthalpies, whether or not N is allowed to float.
- Heat error does not propagate strongly to direct enthalpies, at least for the C values considered here.
- The decision to float or fix N can be guided by examining the discrepancy between direct and van't Hoff enthalpies with both N float and N fixed: a lower discrepancy correlates to a lower error relative to the true enthalpy.

Transparency document

The Transparency document associated with this article can be found, in online version.

Acknowledgements and disclosures

M.K.G. has an equity interest in, and is a cofounder and scientific advisor of VeraChem LLC. S.A.K. acknowledges training support from the Molecular Biophysics training grant (T32 GM008326). This publication was supported by the National Institute of General Medical Sciences of the National Institutes of Health (NIH) Grant GM61300. These findings are solely of the authors and do not necessarily represent the views of the NIH.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.bbagen.2017.11.020.

References

[1] J. Yin, A.T. Fenley, N.M. Henriksen, M.K. Gilson, Toward improved force-field accuracy through sensitivity analysis of host-guest binding thermodynamics, J. Phys. Chem. B 119 (2015) 10145–10155.
[2] J.H. van't (Jacobus H. van't) Hoff, Etudes de dynamique chimique, Frederik Muller, Amsterdam, 1884.
[3] T. Wiseman, S. Williston, J.F. Brandts, L.-N. Lin, Rapid measurement of binding constants and heats of binding using a new titration calorimeter, Anal. Biochem. 179 (1989) 131–137.
[4] H. Naghibi, A. Tamura, J.M. Sturtevant, Significant discrepancies between van't Hoff and calorimetric enthalpies, Proc. Natl. Acad. Sci. U. S. A. 92 (1995) 5597–5599.
[5] Y. Liu, J.M. Sturtevant, Significant discrepancies between van't Hoff and calorimetric enthalpies III, Biophys. Chem. 64 (1997) 121–126.
[6] L.S. Mizoue, J. Tellinghuisen, Calorimetric vs. van't Hoff binding enthalpies from isothermal titration calorimetry: Ba^{2+}-crown ether complexation, Biophys. Chem. 110 (2004) 15–24.
[7] G. Weber, van't Hoff revisited: enthalpy of association of protein subunits, J. Phys. Chem. 99 (1995) 1052–1059.
[8] G. Weber, Persistent confusion of total entropy and chemical system entropy in chemical thermodynamics, Proc. Natl. Acad. Sci. U. S. A. 93 (1996) 7452–7453.
[9] A. Holtzer, Comment on 'van't Hoff revisited: enthalpy of association of protein subunits', J. Phys. Chem. 99 (1995) 13048–13049.
[10] J. Ross, Comments on the article "Persistent confusion of total entropy and chemical system entropy in chemical thermodynamics" [(1996) Proc. Natl. Acad. Sci. USA 93, 7452–7453], Proc. Natl. Acad. Sci. U. S. A. 93 (1996) 14314.
[11] J.R. Horn, D. Russell, E.A. Lewis, K.P. Murphy, van't Hoff and calorimetric enthalpies from isothermal titration calorimetry: are there significant discrepancies? Biochemistry (Mosc) 40 (1774–1778) (2001).
[12] B.M. Castellano, D.K. Eggers, Experimental support for a desolvation energy term in governing equations for binding equilibria, J. Phys. Chem. B 117 (2013) 8180–8188.
[13] H.S. Muddana, A.T. Fenley, D.L. Mobley, M.K. Gilson, The SAMPL4 host-guest blind prediction challenge: an overview, J. Comput. Aided Mol. Des. 28 (2014) 305–317.
[14] A.T. Fenley, N.M. Henriksen, H.S. Muddana, M.K. Gilson, Bridging calorimetry and simulation through precise calculations of cucurbituril–guest binding enthalpies, J. Chem. Theory Comput. 10 (2014) 4069–4078.
[15] N.M. Henriksen, A.T. Fenley, M.K. Gilson, Computational calorimetry: high-precision calculation of host–guest binding thermodynamics, J. Chem. Theory Comput. 11 (2015) 4377–4394.
[16] S.A. Kantonen, N.M. Henriksen, M.K. Gilson, Evaluation and minimization of

uncertainty in ITC binding measurements: heat error, concentration error, saturation, and stoichiometry, Biochim. Biophys. Acta (2016), http://dx.doi.org/10.1016/j.bbagen.2016.09.002.
[17] V.J. LiCata, C.-C. Liu, Chapter nine - analysis of free energy versus temperature curves in protein folding and macromolecular interactions, in: Michael L. Johnson, J. M. H., G. K. A. (Eds.), Methods in Enzymology, 488 Academic Press, 2011, pp. 219–238.
[18] J. Carrazana, A. Jover, F. Meijide, V.H. Soto, J. Vazquez Tato, Complexation of adamantyl compounds by beta-cyclodextrin and monoaminoderivatives, J. Phys. Chem. B 109 (2005) 9719–9726.
[19] J.B. Chaires, Possible origin of differences between van't Hoff and calorimetric enthalpy estimates, Biophys. Chem. 64 (1997) 15–23.
[20] W.B. Turnbull, A.H. Daranas, On the value of c: can low affinity systems be studied by isothermal titration calorimetry? J. Am. Chem. Soc. 125 (2003) 14859–14866.
[21] B.A. Pethica, Misuse of thermodynamics in the interpretation of isothermal titration calorimetry data for ligand binding to proteins, Anal. Biochem. 472 (2015) 21–29.
[22] S.E. Boyce, J. Tellinghuisen, J.D. Chodera, Avoiding accuracy-limiting pitfalls in the study of protein-ligand interactions with isothermal titration calorimetry, bioRxiv (2015) 023796, http://dx.doi.org/10.1101/023796.
[23] Y. Liu, J.M. Sturtevant, Significant discrepancies between van't Hoff and calorimetric enthalpies. II, Protein Sci. Publ. Protein Soc. 4 (1995) 2559–2561.
[24] J. Tellinghuisen, J.D. Chodera, Systematic errors in isothermal titration calorimetry: concentrations and baselines, Anal. Biochem. 414 (2011) 297–299.
[25] D.A. McQuarrie, Statistical Mechanics, Harper & Row, 1973.
[26] M.K. Gilson, J.A. Given, B.L. Bush, J.A. McCammon, The statistical-thermodynamic basis for computation of binding affinities: a critical review, Biophys. J. 72 (1997) 1047–1069.
[27] J.R. Horn, J.F. Brandts, K.P. Murphy, van't Hoff and calorimetric enthalpies II: effects of linked equilibria, Biochemistry (Mosc) 41 (7501–7507) (2002).
[28] S.M. Agena, I.D.L. Bogle, F.L.P. Pessoa, An activity coefficient model for proteins, Biotechnol. Bioeng. 55 (1997) 65–71.

Chapter 5, in full, is a reprint of the material as it appears in “Accounting for apparent deviations between calorimetric and van't Hoff enthalpies”, SM Kantonen, NM Henriksen, MK Gilson. Biochimica et Biophysica Acta (BBA)-General Subjects (2018). The dissertation author was primary investigator and author of this publication.

6 CONCLUSIONS

Molecular dynamics condenses an incredibly complex problem down to an approximate classical form. If the perceived promise of accurately predicting experimental results *in silico* is to be met, advancements will need to continue to be made in refining force-fields. Part of the philosophy of this work is that force-field development is a multi-phase endeavor, and that careful attention needs to be paid to each component of the overall development.

One of the key points this thesis presents is the condensation of the force-field optimization problem. The SDLJ force-field, discussed in Chapter 2, was able to train on just seven small molecules, with two properties per molecule, while reducing the parameter space to just ten parameters. Additionally, we were able to remove atom-types from our method, allowing each molecule to have its own unique set of parameters. By condensing the parameter space down to just ten parameters, we can avoid potential over-fitting while also making the overall optimization problem easier. The SDLJ method also is advantageous is that it incorporates both quantum mechanical data and experimental data into the overall optimization process, providing a more robustly informed set of parameters.

In terms of experimental work, this thesis has provided some insight into single site modification of host molecules and the subsequent analysis of binding via isothermal titration calorimetry. These components are both important parts of force-field development, even though their disciplines lie outside computational chemistry. The ability to reliably modify host molecules is of great use to the computational chemist, as it provides a more diverse set of experimental data to explore computationally. Without a wide range of reliable experimental data to compare computational results to, force-fields

may become focused on a narrow range of compounds, providing inaccurate or misleading results for new or interesting compounds.

The ability to accurately measure the binding thermodynamics is another experimental component that is remarkably important to force field optimization. Noise in the data used for optimization will move the force field parameters in an inaccurate direction, so having some understanding of the nature of errors, and some understanding of how to check for possible inconsistencies, is of great use to the force field developer. Here we showed an extensive analysis of error in isothermal titration calorimetry (Chapter 4), and developed software for fitting ITC data based on this analysis. This software provides more realistic estimates of error than the commonly provided least-squares fitting program used to fit ITC data. Additionally, we showed that van't Hoff enthalpies and calorimetric enthalpies should be in agreement (Chapter 5), and that checking for disagreement between the two can inform our understanding of the accuracy of the data.

6.1 FUTURE DIRECTIONS

The most straightforward thing to test next is how the force field parameters generated from the SDLJ method do when applied to simulations to calculate binding thermodynamics. While transferability exists for neat liquid simulations, it remains to be seen if the same transferability applies when calculating binding free energies. This can be tested on some of the functionalized cyclodextrins our laboratory has generated, but can also be applied to prior existing data on unmodified β-cyclodextrins.

The next step would be to incorporate host-guest binding data into the SDLJ optimization process. We have designed a method, which reduces parameter space, which can now be tested when training on experimental binding free energies. An

important aspect to consider is the water model used. Although an existing water model might be used, it would seem most interesting to apply the SDLJ method to water as well, generating water parameters from electron densities along with the parameters of the host and guest compounds. This would provide a fully self-consistent set of parameters.

There are additional considerations that can be taken in exploring the SDLJ concept. Firstly, we can explore how to incorporate the partitioned charges into the simulation; the current work on SDLJ uses AM1-BCC charges, as outlined in Chapter 2, but the stockholder atoms-in-molecules approached use to obtain the β values also provides partial atomic charges. Another mapping parameter could be incorporated to scale the charges (conserving net charge scaling all at once, i.e., doubling the partial charges on all atoms) and thus optimize the overall level of polarity, much as done in a related project from our lab that is currently under review. Other more complicated modifications can be made, such as to the treatment of electrostatics in the potential function. For example, once β values are obtained for atoms in a given molecule, the electrostatic energy can be rewritten as to account for the overlap of densities (which also are closely related to β). This would allow effects such as charge penetration to be modeled, which could lead to significant improvements in the accuracy of noncovalent binding calculations.